14

Educational Evaluation and Examinations

A TREND REPORT

B.K. PASSI
R.C. HOODA

The University Education Commission, set up in 1948 to investigate the condition of the system of university education, remarks, "We are convinced that if we are to suggest any single reform in university education it would be that of examinations." Related to school education, the Secondary Education Commission (1952) observes, "The examinations today dictate the curriculum instead of following it, prevent any experimentation, hamper the proper treatment of subjects and sound methods of teaching, foster a dull uniformity rather than originality, encourage the average pupil to concentrate too rigidly upon too narrow a field and thus help him to develop wrong values in education. Pupils assess education in terms of success in examinations. In short, external examinations are exercising a restricting influence over the entire field of education to such an extent as almost to nullify its real purpose." Both the Commissions have highlighted the importance of examination system. The Education Commission (1964-66) again emphasized the need for setting up special units for examinations and evaluation which should prepare and implement a programme of reforms in close collaboration with a central examination unit to be set up at the UGC headquarters.

The UGC has played a leading role in improving examination systems at the university stage. It has supported the willing universities in setting up examination research and reform units. It has given substantial grants and support to universities for improving examination in higher education. The document of the UGC entitled Examination Reform — A Plan of Action (1973), says that examinations have dominated the educational process, and external examinations have encouraged selective study and cramming. Examination marks lack in reliability and validity and unfair means in the examination have increased tremendously. The crippling effect of external examinations on the quality of instruction has compelled various agencies to bring about improvement in the present system. In this context, the UGC has proposed the implementation of innovative programmes like internal assessment, grading system, question bank, etc. It was desired that there should be a time-bound programme for the successful implementation of various measures on examination reform in the universities/institutions. The main emphasis has been on the following aspects of examination reforms: (i) continuous sessional evaluation as a supplement to the present final examination, (ii) the development of question banks in order to eliminate some of the shortcomings of setting up examination papers and as a means for revision and modernization of courses of study, (iii) introduction of grade system instead of present marking system in order to increase reliability and bring about better comparability among different subjects, and (iv) introduction of the semester system in order to have greater flexibility.

The Association of Indian Universities (AIU) has given a significant lead to improve the examination system in higher education. In the seventies, the AIU conducted a large number of seminars, developed question banks, published monographs related to various aspects of examinations, publicized different innovations, and conducted a few research studies. As a matter of fact, it can be proudly observed that with limited resources, the

AIU has made a dent in the field of examinations. Very recently the UPSC, the State Public Service Commissions, the NIBM and other similar bodies and institutions have produced indirect impact upon the examination system.

The NCERT has focussed its attention on research and improvement in the examination system at the school stage. The work in this area comprises research, development of materials and operational strategies, training of both pre-service and in-service teachers, publications, and clearing house functions both at the national and the international levels. In addition to this, the NCERT collaborates with the State-level educational agencies like the State Departments of Education, the State Boards of Secondary Education, the State Institutes of Education, the State Councils of Educational Research and Training (SCERT), State Evaluation Units and teacher training institutes both at the elementary and the secondary levels. Some Boards of Secondary Education and SCERT's have also been conducting researches and implementing various innovative programmes in the examination system at the school stage.

Looking to the research in examinations in India, one finds that only a few institutions have been undertaking sustained studies in the area of examinations and evaluation during the last twenty years. Buch and Lele at the M.S. University of Baroda, Gayen at the I.I.T., Kharagpur, Harper at the Ewing Christian College, Allahabad, Taylor at Gauhati University and Bokil at the Maharashtra State Board of Secondary Education have undertaken commendable studies. The NCERT has financed twenty-four research studies, whereas the UGC has financed ten research studies in this area. The NCERT has directly undertaken more than ten research studies.

The studies in the area of evaluation and examinations have been compiled and/or reviewed by Long and Mehta (1966), Pareek and Kumar (1966), Pareek and Sood (1971), Dave (1968), Buch (1972), Passi and Padma (1974), and Passi and Sansanwal (1979). A synoptic overview of the nature of these works would help to take stock of the research completed in the area of educational evaluation, examinations and achievement tests.

In the First Mental Measurement Handbook for India, Long and Mehta (1966) included 326 tests. Of these, one hundred were classified under intelligence, ninety-six under achievement, sixty under aptitude, fifteen under interest, forty-five under personality and ten under the miscellaneous category. Out of the ninety-six tests under achievement, thirty-five were related to mathematics, thirty to languages, eighteen to sciences, and thirteen to social studies. Pareek and Kumar (1966), in their Directory of Behavioural Science Research in India from 1925 to 1965, covered 902 studies in the area of achievement testing. Of these, there were 173 on examination, 138 on general achievement, 104 on languages, 133 on mathematics, 103 on social studies, 130 on sciences, 28 on other subjects and 93 on scholastic backwardness. It may be noted that Pareek and Kumar (1966) referred to the studies related to the fields of psychology, sociology, cultural anthropology, education, social work, community development, agricultural extension and political behaviour. In continuation to Pareek and Kumar (1966), Pareek and Sood (1971), in the Directory of Indian Behavioural Science Research, covered 326 studies related to achievement testing. The break-up of these studies is as follows: examination — 73, general achievement — 44, language — 45, mathematics — 42, social studies — 42, science — 23, other subjects — 24 and scholastic backwardness — 33.

Dave (1968), in the Third Indian Year-book of Education — Educational Research — has reviewed 407 studies from 1941 to 1966. These studies cover individual research papers published in journals, aided and unaided research projects, M.Ed. dissertations and Ph.D. studies. It is important to mention here that the review is mostly based upon either the M.Ed. dissertations or the published work. Mitra (1968) reviewed tests and measurement in the Third Indian Year-book of Education. Buch (1972) reviewed published research studies in the area of educational psychology. Within this broad area, achievement tests and evaluation could not be given significant space and importance. Two very comprehensive surveys in Educational Research in India have been completed by Buch. In A Survey of Research in Education, Passi and Padma (1974) developed a trend report based on ninety-five doctoral and institutional research projects. Again, Passi and Sansanwal (1979) wrote a trend report on Educational Evaluation and Examinations based on almost all the research studies conducted at doctoral and institutional levels in India in the Second Survey of Research in Education.

The present trend report is based on 198 studies available in the area of examinations and achievement testing both at the Ph.D. and the project levels. These studies have been classified into six different areas which can be distinguished in terms of emphasis only. The six areas are: (i) Achievement Tests, (ii) Diagnostic Tests, (iii) Examinations, (iv) Factors Affecting Achievement, (v) Prediction-Admission-Promotion, and (vi) Failures. The subsequent discussion in this trend report follows this sequence.

Nature of Studies

The present review is based upon 198 studies: 80 for Ph.D. in Education, 10 for Ph.D. in Psychology and 108 research projects. This review refers to the research studies completed as early as 1943 and as late as March 1983.

A detailed picture of the six sub-areas of the problems of research in educational evaluation and examinations and the developmental trends is given in Table 1.

TABLE 1
AREA-WISE AND PERIOD-WISE DISTRIBUTION OF STUDIES

	Up to 1949	*1950-54*	*1955-59*	*1960-64*	*1965-69*	*1970-74*	*1975-79*	*1980*	*Total*
Achievement Test	—	—	2	10	16	17	16	8	69
Diagnostic Test	—	—	—	—	5	5	2	2	14
Examinations	—	—	2	11	12	13	15	9	62
Factors Affecting Achievement	1	2	5	2	4	3	2	1	20
Prediction-Admission-Promotion	—	2	2	6	4	3	3	1	21
Failures	—	—	2	4	2	—	3	1	12
Total	1	4	13	33	43	41	41	22	198

Table 1 includes data indicating period-wise growth of research work in the sub-areas of Educational Evaluation and Examinations. It may be noticed here that up to 1954, very little research work was done and it was restricted to two sub-areas, viz. Factors affecting Achievement and Prediction-Admission-Promotion. In the years 1960 to 1974, there was a heavy concentration (about three-fourths) in the sub-areas of Achievement Tests and Examinations. No research in the area of Diagnostic Tests was done up to 1964. The studies related to Failures were undertaken only after 1954, and that, too, only at the institutional level. This suggests that there existed a trend in the nature of research problems undertaken over the years.

A closer scrutiny of the Ph.D. research in education and other disciplines as well as institutional and individual projects indicates that only one research study was completed as the Ph.D. level before Independence. After Independence, there was a gradual increase in Ph.D. and institutional research in this area. During the period 1965-69, forty-three studies were completed. Another interesting feature is that out of the 198 studies during the period 1944-82, 158 studies were completed during the period 1960-79. Achievement Tests claim 69 studies and Examinations 62 studies. The researchers, it seems, are not attracted to the study of the problems of failures where only twelve studies are cited. Even these twelve studies are institutional projects rather than Ph.D. research.

Methodology

Sample

Different investigators aiming at standardizing tests have used samples of different sizes and natures. About two-thirds of these tests have been standardized on samples of less than three thousand students and there are only a few such test standardization studies which have used samples of more than seven thousand students. The samples from the school stage have been drawn from all the grades. In some of the studies which aim at constructing entrance tests, samples have been drawn from the respective university faculties like arts, science, medicine, education, etc. Depending upon the location of the university, institution, investigator, purpose of the study and so on, samples have been drawn from different parts of the country. The methodology for drawing the samples happens to be simple randomization, multi-stage randomization, stratified sampling, random stratified, clustered, etc., depending upon the nature of the related population and the purpose of the study. The variations about the sample, the size and the methodology of sampling are very wide.

Tool Construction

Most of the investigators have developed achievement tests or used examination marks for conducting their studies. The methodology employed for developing the achievement tests and question papers for the examination system includes the usual standard steps. The test constructor plans, prepares, tries out, and establishes norms. During planning, the content analysis, analysis of instructional objectives, preparation of blue print, nature of items, and the length of the test are decided upon. For the preparation of the first draft, the investigator constructs the items, edits them, and discusses the same with the experts and the practitioners. In the last two phases of test construction, he carries out item analysis, and establishes reliability, validity and norms for the test.

Item analysis. While going through the procedure employed in item analysis in the set of studies given in this report, one finds that two item indices, namely, item difficulty and item discrimination, have been widely used. One also observes that researchers have not explicitly stated the nature of sample for item analysis. It must be noticed that the item analysis sample should be

representative of the population for which the test is designed. If a test comprises subtests, it is usually more appropriate to perform separate item analysis within each subtest. If the test performance is dependent upon sex, caste, creed, religion, and locality of the responders, separate item analyses be carried on for these subgroups. Of course this also holds true for reliability and validity studies.

Reliability. The researchers have followed the usual methods of establishing reliability, such as test-retest, parallel form, split-half, and internal consistency. On the basis of review of the studies included in the area of Educational Evaluation and Examinations, it has been observed that the time gap between test-retest administrations had been fixed casually. None of the studies has introduced multiple time gaps. The multiple time gaps with test-retest or parallel forms should be studied to establish the stability coefficients. It is also observed that in most of the studies related to essay-type tests, the content reliability has not been computed. The marks reliability for essay-type tests can be obtained as the product of content reliability and examiner reliability. In essay-type tests, the marks reliability is the content reliability attentuated by examiner unreliability. It may also be noted that the reliability of a test becomes more meaningful when the standard error of measurement is reported. Surprisingly, very few studies have reported this. It is suggested that large-scale multifacet studies of reliability aiming to find out the error components due to items, scorers, instructions, occasions, examinee, etc., be conducted.

Validity. Content validity, concurrent validity, and factorial validity have been established in different studies. The predictive validity is almost missing. The usual criterion variables of establishing concurrent validity are teachers' rating, scores on the teacher-made tests, school examination marks, board examination marks, intelligence tests, and some other standardized tests. Some experts believe that such criterion measures suffer from a few defects. They claim that the desired qualities of a criterion are: relevance, freedom from bias, reliability, and availability. Researchers need to explicitly demonstrate the presence of such qualities in the chosen criterion for validation. The present researches do not indicate the factors influencing the validity of their tools.

Norms. A few test constructors have established norms in the form of percentiles, standard scores, and grade norms. The absence of comprehensive norms in standardized tests will restrict their use. As a matter of fact, criterion reference testing has important bearing upon educational achievement. No study has developed criterion reference tests. In some of the achievement tests the investigators have not either developed the norms or have developed local norms. Perhaps, it suited their purpose. But, in order to develop national achievement tests related to all-India curricula, a few investigations have to be undertaken.

Statistical Designs

From the methodological point of view, the break-up of 198 studies is as follows: (i) descriptive and correlational — 163, (ii) factor analytical — 17, (iii) regression — 14, (iv) experimental — 3, and (v) regression and factor analytical — 1. The three experimental studies have used the simple pretest and posttest design with one treatment and one control group. The factor analytical studies have aimed at either seeing the factorial validity of different instruments or classifying different school subjects into new families of subjects or seeing the nature of factors involved in the test batteries. The regression studies have aimed at establishing multiple regression equations and multiple correlation between the predictors like the S.S.C. examination marks and criteria, like college grades.

ACHIEVEMENT TESTS

The review of studies on achievement tests is presented in the order of general scholastic achievement tests, achievement tests in languages, achievement tests in social sciences, achievement tests in mathematics, achievement tests in sciences and, finally, some miscellaneous tests.

General Scholastic Achievement Tests

Standardized general scholastic achievement tests are frequently used in the form of test batteries. There are always variations in subject matter content from one curriculum to another and also differences in grade placement of instructional materials from state to state. It is very unlikely that the various sections of a test battery will be uniformly applicable to the instructional programme of any given school. This limitation is especially pronounced in content-oriented test batteries. It is of less significance in batteries designed to measure basic skills and general educational development. The tests of general scholastic achievement are tests of general educational development. They are of special interest because they measure complex learning outcomes that cut across subject matter lines and are common to the major content areas of the school. They emphasize understanding, interpretative skills, and the ability to apply

knowledge and skills to new situations. Since the learning outcomes measured by these tests are closely related to the ultimate objectives of education, such tests are especially likely to have a desirable influence on the curriculum and the teaching methods of the institution. Tests of generalized abilities help teachers to gauge the general rate of student learning and are useful for grouping students for various educational tasks. Keeping in view the usefulness of general scholastic achievement tests for teachers and guidance workers, concentrated efforts are needed for their development. Some efforts have been made by Parikh (1946 b), Liddle (1965), Jha (1974), Sharma (1975), Sharma (1976), Patel (1977), De (1979), and Shah (1982). Lele and Parikh (1965) prepared a Scholastic Aptitude Test for admission to preparatory science courses. The test consisted of three subtests in English, numerical ability, and abstract reasoning. Liddle (1965) standardized an Academic Aptitude Test (AAT) for high school students of Uttar Pradesh. It included subtests of vocabulary, numerical computation, sentence completion and mathematical reasoning. The reliability coefficients for each subtest as well as the total test ranged from 0.83 to 0.89. The concurrent validity coefficients against scholastic achievement it terms of total scores ranged from 0.46 to 0.76. Jha (1974) developed a battery of tests for measuring the basic skills related to arithmetic, geometry, and languages. Sharma (1975) developed a biological science aptitude test for school students. Sharma (1976), developed a battery of tests in general science and mathematics for the delta class pupils studying in different parts of Rajasthan. Patel (1977) developed scholastic achievement tests in Hindi, history and geography for Classes V, VI and VII for the pupils studying in the State of Gujarat. De (1979) developed a scholastic achievement test in geography for the secondary students of West Bengal. It included subtests of classification, analogy, information, reasoning, comprehension, computation skill, sentence completion, analysis of relationship and map identification. The reliability coefficients for each subtest as well as the total test ranged from 0.45 to 0.93. Shah (1982) developed a Primary School Achievement Test (PSAT) for the pupils of Grade VII in the State of Gujarat. The PSAT consisted of four subtests namely, vocabulary, routine computation, sentence completion, and mathematical reasoning. The reliability coefficients by different methods ranged from 0.83 to 0.87.

The general scholastic achievement tests have been developed covering the subjects, English, Hindi, general science, mathematics, history and geography. These tests were constructed for the students of Grades V to XI, in the States of Uttar Pradesh, Gujarat, Punjab, Rajasthan, and West Bengal. They have not been developed for the subjects of social studies, economics, physics, chemistry, civics, regional languages, etc. The general scholastic achievement tests for the affective and psychomotor domains have not been developed at all. Some of the tests developed were content-oriented rather than oriented to basic skills and general educational development. Efforts need to be made to develop tests related to basic skills and general educational development, so that they could be used for a variety of purposes.

Achievement Tests in Languages

It is generally accepted that there are four major skills which can be attained in different degrees in learning a language. These skills are: listening, speaking, reading and writing. Adequate and valid measures of proficiency in language skills are needed both in instruction and research. Most standardized tests of languages assess competence only in relative terms – percentiles or other derived measures based upon the performance of typical groups in typical courses of instruction. These measures do not provide direct information concerning the absolute levels of success. One finds from the research literature that language-skill-based achievement tests have been developed in English, Hindi, Gujarati, Oriya, Marathi and Kannada.

In the case of achievement tests for English there are only eleven tests for Grades VI to XI. Aram, Rangaswamy and Feroze (1957) standardized an achievement test in English for the middle stage students of Coimbatore district of Tamil Nadu. The test included subtests related to language usage, spellings, punctuation, capitalization, reading, comprehension and vocabulary. Buch, Patel and Kotwal (1960) standardized achievement tests in English for Classes VIII, IX and X. The test-retest reliability coefficients ranged between 0.88 and 0.96. Misra (1970) and Deshpande (1972) made similar efforts for students in Assam and Maharashtra, respectively. Sinha (1967) attempted to explain the factorial structure of different aspects of attainment in English by constructing subtests related to formal grammar, punctuation, spelling, translation, syntax, expression, legibility, pronunciation, vocabulary, comprehension, and applied grammar. He found and named a factor, Automization Factor, having high loadings on pronunciation, spelling and expression. The second factor had high loadings on comprehension, vocabulary, and pronunciation. The third factor shared loadings with

formal grammar, punctuation, and applied grammar. Chatterji, *et al.* (1970) developed a Test of English Knowledge and Comprehension at the higher secondary level. It covered English usage, word meaning, grammar, spelling, and comprehension. The reliability coefficients estimated for different subtests by using K-R formula 21 ranged between 0.53 and 0.81. Singh (1978) constructed a battery of objective tests for the assessment of proficiency in writing English composition. The battery consisted of seven tests on spelling, punctuation, vocabulary (phrases), vocabulary (words), paragraph organization, applied grammar, and handwriting. The handwriting scale was evolved by Thurstone's equal appearing interval method.

Patel (1971) standardized a Silent Reading Comprehension Test in English for the S.S.C. pupils of Gujarat. The test contents were related to ability to note significant details, to select appropriate meaning of a word in content, to read maps and tables, to follow the sequence of events, to draw generalization, etc. The reliability coefficients worked out by different methods ranged from 0.91 to 0.96. The concurrent validity coefficient against the criterion of teachers' opinion about pupils' comprehension was 0.47. The tests shared loadings with three factors, viz., (i) Ability to Grasp, (ii) Word Meaning, and (iii) Perceptual Speed Factor. Skariah (1981) standardized an Oral Reading Comprehension Test in English for the pupils of Class VIII of Gujarat. Keeping in view the readability of the passages and discriminative value and difficulty value of test items, eight passages were included in the final form of the test. The final form consisted of nine subtests including one for practice. The reliability coefficients of the test were worked out by employing the methods of test-retest, split-half, and K-R formula. These coefficients ranged between 0.82 and 0.97. The concurrent and predictive validity coefficients were established by correlating the scores on the Oral Reading Comprehension Test with the teachers' opinions and the marks obtained in the examination.

Related to Hindi as language, a few achievement tests have been constructed by Shukla and Tutoo (1959), the CIE (1962), the Gujarat Research Society (1963), Jha *et al.* (1964), Sharma (1967), Deshpande (1972), Gaur (1973), Giri (1976), Verma (1977), and Joshi (1980). These studies provided tests for Grades V to S.S.C. level. It is to be noted that achievement test construction in Hindi is localized in Delhi, Haryana, Gujarat, and Maharashtra only. Shukla and Tutoo (1959) constructed an achievement test in Hindi for Grade VI employing Delhi sample. CIE (1962) constructed a test in Hindi for Grade VIII students of Delhi. The test-retest reliability coefficient was 0.89, and the Spearman-Brown reliability coefficients for girls and boys were 0.93 and 0.95, respectively. The concurrent validity coefficient against a teacher-made test was 0.69. The Gujarat Research Society (1963) developed a test of Hindi for Gujarati-medium students of Standards V, VI and VII of Bombay. Jha, *et al.* (1964) developed achievement tests in Hindi for the students of Grades V to VIII. The subtests included were vocabulary, comprehension, recognition of tenses, spellings, sentence structures, etc. The reliability coefficient by K-R formula 20 was reported to be 0.97. The test was validated and norms were established. Sharma (1967) developed a test for the students of Grades VIII to X, whereas Deshpande (1972) made similar efforts for the higher secondary school students of Maharashtra. Gaur (1973) standardized achievement tests in six different aspects of Hindi, viz., spelling, vocabulary, applied grammar, formal grammar, idioms and proverbs, and comprehension. The test was standardized for the students of the matriculation level of Haryana. The split-half reliability coefficients for the tests varied from 0.83 to 0.92. Giri (1976) standardized a Test of Hindi Vocabulary for Classes VI to IX. The coefficient of reliability calculated by K-R formula was 0.98. The concurrent validity coefficients were calculated by correlating the test scores with the school examination marks and Mehta's Intelligence Test. The validity coefficients ranged from 0.55 to 0.89. Verma (1977) developed an Achievement Test in Hindi for Classes VI and VII covering written language abilities, viz., word meaning, language uses, grammatical form recognition, sentences, analogy of words, and comprehension. The test-retest reliability coefficients varied from 0.81 to 0.84 and the split-half reliability coefficients varied between 0.72 and 0.88. The validity coefficients found against the annual examination scores were 0.72 for Class VI and 0.69 for Class VII. Joshi (1980) constructed and standardized a Proficiency Test in second language (Hindi) for the senior primary school teachers of Gujarat. The test covers three aspects of Hindi, viz., vocabulary, written expression, and reading comprehension. The reliability coefficients were worked out by employing the methods of test-retest, split-half, K-R formula, and Rulon formula. The reliability coefficients ranged from 0.82 to 0.98.

Amongst regional languages, Gujarati got the maximum attention from the researchers of the country in the area of achievement test construction. Efforts in this direction have been made by Buch, Patel and Kotwal (1960), Bhagatwala (1960), Maniar (1961), the

Gujarat Research Society (1963), Bhatt (1971), Krishnamurti (1971), Maniar (1973), Pandya (1973), Parekh (1973), Desai (1974), Gokil (1974), Modi (1975), Bisnagari (1976), Patel (1978), and Upadhyaya (1979). Achievement tests in Gujarati are available for all grades from V to pre-university and for children in the age group three to five years. Buch *et al.* (1960) constructed achievement tests in Gujarati for the students of Grades VIII, IX and X. Rational equivalence reliability coefficient, content validity, concurrent validity, norms, etc., have been worked out. Maniar (1961) made efforts to develop a vocabulary test in Gujarati for the children of Bombay in the age range thirteen to seventeen. The Gujarat Research Society (1963) developed achievement tests in Gujarati for the children of Bombay of Standards V to VII. Bhatt (1971) developed an achievement test in Gujarati for the students of Grade VIII and prepared the norms for Gujarat. By employing test-retest, split-half, Rulon's formula, and K-R formula 20, reliability coefficients were worked out, which ranged from 0.89 to 0.97. The concurrent validity against the school examination marks in Gujarati was found to be 0.96. Pandya (1973) standardized a Language Ability Test in Gujarati for the college entrants of Gujarat. The test consisted of four parts, viz., vocabulary, sentence structure, spelling and punctuation, and comprehension. The test-retest reliability was found to be 0.82. The concurrent validity against the marks in Gujarati language at the S.S.C. Examination was 0.75. Desai (1974) standardized Language Development Tests for Gujarati children in the age group three to five. The test includes items related to comprehension, sound discrimination, articulation, and oral expression. Bisnagari (1976) standardized a Gujarati Handwriting Scale for the pupils studying in Standards V to XI in the schools of Gujarat. Upadhyaya (1979) standardized a Proficiency Test in first language for the primary school teachers of Gujarat. The test has three subtests, viz., vocabulary, written expression, and reading comprehension. The reliability coefficient was 0.82 by test-retest, 0.98 by analysis of variance, 0.90 by split-half, and 0.94 by formula K-R 20. Bhagatwala (1960) standardized a Silent Reading Test in Gujarati for Grades VIII to XI. The test intended to measure the speed of reading, word meaning, and comprehension. Krishnamurti (1971), at Madras, also developed a Reading Readiness Test for School Children. Maniar (1973) developed a Silent Reading Test in Gujarati for the pupils of Standard VIII in Gujarat. The test measures reading rate, reading comprehension, vocabulary, location of information, etc. Various types of reliability coefficients ranged from 0.69 to 0.98. Parekh (1973) standardized a Silent Reading Test in Gujarati for pupils studying in Standard IX. The test includes nine subtests related to different aspects of silent reading. Different types of reliability coefficients ranged from 0.68 to 0.97. Gokil (1974) standardized a Silent Reading Test in Gujarati for pupils of Standard VII. The test consisted of seven subtests. The reliability coefficients were worked out by various methods and they ranged from 0.89 to 0.96. The only research project related to standardization of Listening Comprehension Test in Gujarati was completed by Modi (1975). The test had five types of items related to listening in different situations, attention, memory, auditory resistance, and reasoning. It was developed for Gujarati children of Standard VIII. The test-retest reliability coefficient was 0.80. Patel (1978) standardized a Listening Comprehension Test in Gujarati for pupils of Classes VIII, IX and X, to study the effect of exercises for improving it. The improvement programme consisted of three types of exercises viz., listening games, developing listening through the use of the textbook, and listening comprehension exercises based on selected listening material. The developed programmes were found to be effective and could be used to improve the listening comprehension of pupils.

There are three other studies related to achievement test construction in three other regional languages — Kannada, Oriya, and Marathi. Dash (1967) standardized an achievement test in Oriya for Grade VII students, whereas Deshpande (1972) developed an objective assessment tool in Marathi for the students appearing for the secondary education examination in Maharashtra. Shivananda (1981) standardized Reading Tests in Kannada for pupils of Standards V to VII, separately. The test for each standard was developed covering the aspects of vocabulary, opposites, word discrimination, reading comprehension, and sentence formation. The final form of the test comprised thirty items for each of the five subtests. The reliability coefficients worked out by different methods ranged from 0.58 to 0.96, from 0.58 to 0.97 and from 0.60 to 0.96 for Standards V, VI and VII, respectively. The validity coefficients ranged from 0.60 to 0.70, 0.60 to 0.69 and 0.50 to 0.56 for Standards V, VI and VII, respectively.

To sum up, achievement test construction in languages at the national and the regional levels needs more extensive and concentrated efforts. Except in four regional languages, namely, Gujarati, Oriya, Kannada, and Marathi, no systematic and sustained attempts have been made to standardize tests in regional languages at the Ph.D. level or at the institutional level. It is interest-

ing to see that the studies in the areas of reading speed, reading comprehension, and listening comprehension are restricted to the Gujarati language. All studies have been localized in Gujarati except one study by Maniar (1961). Research studies at the M.Ed. level aiming at measuring achievement in regional languages should be coordinated in order that each regional language has achievement tests for all grade levels.

Achievement Tests in Social Sciences

An effective programme of social studies is essential for the development of good citizenship. Some experts believe that an integrated and coordinated programme of social studies centring round the study of men and his environment should be followed at the lower primary stage. At the elementary stage, pupils should be gradually introduced to an appreciation of history, geography, and civics as separate subjects. In order to increase the effectiveness of the social studies programme, a sound achievement testing service is essential. The present inadequacy of tests in social sciences urgently demands the designing of tests in this important area. Some efforts for achievement test construction in the area of social sciences were made by Aram *et al.* (1957), Shukla and Tutoo (1959), Buch *et al.* (1960), the Gujarat Research Society (1963), Saraf (1964), the SIE, Kerala (1965), Dash (1967), Muzaffar (1967), Srivastava (1967), Misra (1968), Vanajakshi (1970), Misra (1970), Deshpande (1972), and Sharma (1981). They developed tests in social studies, history, geography, and civics only. The social studies tests were constructed for students of Grades IV to VIII, in history for Grades V to XI, in geography for Grades V to VIII, X, and XI, and civics for Grades IX to XI.

Aram *et al.* (1957) developed an achievement test in social sciences for the middle school students of Coimbatore district in Tamil Nadu. The SIE, Kerala (1965) standardized an achievement test in social studies for the students of Standard VIII in Kerala. The test had a split-half reliability of 0.72. Similarly, Vanajakshi (1970) developed a test for the Grade VII students of Andhra Pradesh. The test-retest, split-half, and K-R formula 20 reliability coefficients were 0.91, 0.94 and 0.90, respectively. Again, for the students of Grade VII in Orissa, an achievement test in social studies was standardized by Dash (1967).

Buch *et al.* (1960), in their study, Baroda Mental Measurement Series, reported the construction of achievement tests in history for students of Grades VIII, IX and X. The Gujarat Research Society (1963) standardized tests in history for students of Grades V to VII. Misra (1968) standardized an achievement test in history for the high school students of Uttar Pradesh. The contents of the test cover four periods of history: Ancient, Pre-Moghul, Moghul and Modern. The test-retest reliability coefficient for the total test was 0.72. For concurrent validity, the criteria of school examination marks and the teachers' estimates of students' achievement in history were used. For the students of Grade XI in Delhi, an achievement test in Indian history was standardized by Saraf (1964). The reliability coefficients by split-half, K-R formula 20, and Guttman's method ranged between 0.79 and 0.85. The concurrent validity coefficients against the criteria of the school board examination marks, teacher ratings, etc., ranged between 0.28 and 0.88. An attempt in this direction was also made by Misra (1970) for the students of Grade X in Assam. Sharma (1981) standardized an Achievement Test in General Knowledge (history and geography) for Class X in Jammu province. The reliability coefficients by split-half, K-R formula 21 and Flanagan's formula were 0.96, 0.84 and 0.76, respectively.

For geography at the school stage, Buch *et al.* (1960), Misra (1970), and Deshpande (1972) developed achievement tests for Grades VIII to X in Gujarat, for Grade X in Assam, and for grade XI in Maharashtra, respectively. Achievement tests in geography were also standardized for Grades V to VII by the Gujarat Research Society (1963). For achievement tests in civics at the school stage, attempts have been made by Srivastava (1967) for the high school students of Uttar Pradesh. Muzaffar (1967) standardized a Battery of Achievement Tests in Civics for the higher secondary stage in Madhya Pradesh. The battery consisted of three tests each for Classes IX, X and XI. The split-half and the Kuder-Richardson formula reliability coefficients ranged between 0.91 and 0.99. The empirical validity of the tests ranged from 0.49 to 0.54.

It is observed that, by and large, achievement tests in social studies, history and geography are available for Grades V to XI employing samples from various States. But, at the same time, three observations can be made regarding achievement tests, namely, (i) they have not been developed for all subjects in social sciences even within a State; (ii) they are not available in all the States even for one subject; and (iii) for certain subjects like economics, sociology, etc., they have not been standardized at all.

Achievement Tests in Mathematics

Unlike languages, social studies and science subjects,

there are relatively more tests constructed in mathematics — arithmetic, algebra, geometry, and trigonometry. Within this area, there are more efforts in arithmetic followed by mathematics as a whole, geometry, and algebra. Tests are available for the students of Grades III to XI in various branches of mathematics. Institutions and universities scattered over Andhra Pradesh, Assam, Bombay, Delhi, Gujarat, Kerala, Maharashtra, and Tamil Nadu have worked in this field.

Different investigators like Aram *et al.* (1957), Maniar (1961), the SIE, Kerala (1965), Dash (1967), Kulkarni *et al.* (1970), Misra (1970), Vanajakshi (1970), and Bhatt (1971) constructed achievement tests in mathematics either for Ph.D. or for institutional projects. Aram *et al.* (1957) prepared an achievement test in mathematics for the middle grade students of Coimbatore district. The test-retest reliability coefficient was 0.89 and the concurrent validity coefficient against a teacher-made test was 0.48. Maniar (1961) constructed the Test of Mathematical Ability. Strictly speaking, it is not an achievement test. It had a positive relationship with achievement in mathematics. The test contents related to the application of four fundamental rules of arithmetic—arithmetical reasoning; fractions, squares, and equations in algebra; geometrical propositions; and spatial relations. The split-half and the parallel form reliability coefficients were 0.98 and 0.90, respectively. The criteria used for validation were achievement in mathematics and the Desai's Group Test of Intelligence, resulting in the corresponding coefficients of 0.48 and 0.62. The SIE, Kerala (1965) standardized an achievement test in general mathematics for the Standard VIII students in Kerala. The test representing items from the areas of skills, concepts, and application had split-half reliability coefficient of 0.88 and concurrent validity coefficient of 0.67 against achievement scores. Dash (1967), while standardizing an achievement test battery for the Grade VIII students of Orissa, developed an achievement test in mathematics which had high loadings on the centroid mathematical factor. Vanajakshi (1970), in Andhra Pradesh, standardized an achievement test in elementary mathematics for students of Grade VII. The test had stability coefficient of 0.90. Misra (1970) developed an objective-type achievement test in mathematics for the high school students of Assam. Bhatt (1971) standardized an achievement test in mathematics for the Grade VIII students of Gujarat. The test-retest, split-half, K-R formula 20, and Rulon's methods resulted in a median reliability coefficient of 0.99, whereas the validity coefficient against the external criterion of school examination marks was 0.87. A very important project in the direction of achievement test construction was carried out by Kulkarni *et al.* (1970). They developed achievement tests in mathematics in all Indian languages at three levels, i.e. primary, middle and high school. These tests were prepared on the lines of Sequential Tests of Educational Progress of Educational Testing Service. The parallel test reliability coefficients ranged from 0.64 to 0.89 for various levels, whereas K-R formula 20 and K-R formula 21 coefficients ranged from 0.71 to 0.88. The predictive validity coefficient against the school board marks was found to be 0.42.

In the area of test construction in arithmetic, Chickermane (1943), Dave (1958), Buch *et al.* (1960), Pendharkar (1965) and Basu (1969) made efforts and provided tests for Grades III to X for the States of Maharashtra, Gujarat, Mysore and West Bengal. Chickermane (1943) developed a test measuring arithmetical ability. Actually it should not be treated as a test falling in the present category of achievement tests in arithmetic. The same is true in the case of Basu (1969) who standardized a group test of problem-solving ability in arithmetic. Dave (1958) constructed achievement tests in arithmetic for the students of different grades in Bombay. The test-retest reliability coefficient was 0.985 and the validity coefficient against the achievement in school examination was 0.44. Buch *et al.* (1960) developed achievement tests in arithmetic for students of Grades VIII to X with K-R reliability coefficients ranging from 0.88 to 0.96 for the different tests in the battery. The Gujarat Research Society (1963) and Pendharkar (1965) standardized tests for the students of Grades V, VI and VII in Greater Bombay. Jha (1974) developed tests related to arithmetic concepts and certain concepts of geometry for measuring the basic skills of students of Classess VI and VII. For algebra, Buch *et al.* (1960) constructed achievement tests for Grades VIII, IX and X. In the case of geometry, Buch *et al.* (1960) constructed achievement tests for Grades VIII to X. Gokhale (1954) developed an achievement test for geometry; Gupta (1974) for algebra, trigonometry and geometry; and Tewari (1975) for arithmetic, algebra and geometry for the purposes of factorial analysis of students' attainment. Sharma (1978) constructed a battery of sequential achievement tests for Classes V to X. The tests were confined to the areas of arithmetic and algebra of school mathematics. The reliability coefficients worked out by the method of rational equivalence ranged from 0.66 to 0.75. The predictive validity coefficients against the school annual examination marks varied from 0.43 to 0.90. Ketkar (1982) standardized Unit Achievement Tests for Standard VIII for the pupils studying in the

Marathi-medium schools of Maharashtra. The number of unit tests was seven (four in algebra and three in geometry). The reliability coefficient worked out by K-R formula 21 ranged from 0.77 to 0.88. The validity coefficients against the school terminal examination ranged from 0.36 to 0.84. Chauhan (1982) constructed and standardized an achievement test in algebra for the Class IX students in Delhi schools with reference to categories of Guilford's Structure of Intellect Model. The test has several subtests, namely, substitution test, inequation test, expression rewriting test, coefficient comparison test, equivalent equation test, and sufficient data test. The eight S-I tests were; letter series (CSS), correlate completion (NSR), word transformation (NST), word relations (CSR), synonyms (CMU), operations sequence (NSS), word patterns (CSI) and necessary facts (CMS). He found that each of the achievement tests in algebra was significantly correlated with each of the S-I tests. Factor analysis of tests of algebra indicated the existence of only one common factor, namely, General Algebraic Reasoning, and factor analysis of S-I tests also indicated only one common factor, namely, Education of Correlates.

It is observed that with the change of curriculum in mathematics the available achievement tests will be outdated. The introduction of new mathematics, statistics, trigonometry, differential calculus, etc., at the school stage, makes it essential for the researcher to develop corresponding measuring tools. Looking into the validity study of achievement tests in mathematics it appears that researchers have used the achievement scores as the external criterion of validation. There is need for establishing other types of validity coefficients.

Achievement Tests in Science

Next to the achievement tests in school subjects of mathematics and languages, researchers have focussed their attention upon the area of achievement test construction in science subjects. Within the family of science subjects there happen to be many studies of test construction in general science, physics, chemistry, botany, zoology and home science.

The studies related to test construction in general science were conducted by Aram *et al.* (1957), Buch *et al.* (1960), Saxena (1960), Gupta (1962), the Gujarat Research Society (1963), the SIE, Kerala (1965), Dash (1967), Sheth (1967), Rup Prakash (1968), Vanajakshi (1970), the SCERT, Hyderabad (1971), Bhatt (1971) and Hira Devi (1973). In the case of general science, tests are available for Grades V to VIII of the States of Tamil Nadu, Punjab, Haryana, Maharashtra, Andhra Pradesh, Gujarat, Kerala and Orissa. Aram *et al.* (1957) developed parallel forms of achievement tests in general science for the middle school students of Coimbatore district. The Gujarat Research Society (1963) and Hira Devi (1973) constructed achievement tests in general science for Grades V, VI, and VII drawing samples from Bombay. Buch *et al.* (1960) in Gujarat, Saxena (1960) and Gupta (1962) in Uttar Pradesh, the SIE (1965) in Kerala and Rup Prakash (1968) in Punjab and Haryana developed achievement tests in general science for Grade VIII. In these studies, the usual methods of finding reliability coefficients were used and the coefficients were, in general, around 0.90. Mostly concurrent validity coefficients have been worked out against the external criterion of marks in school examinations. Buch *et al.* (1960) constructed tests in general science for the students of Grades VIII to X of Gujarat. Dash (1967) developed an achievement test in general science for Grade VII of Orissa. Sheth (1967) and Hira Devi (1973) standardized achievement tests in general science for the students of Grades V, VI, and VII in Uttar Pradesh and Bombay respectively. The reliability coefficients calculated by the split-half, Rulon's, and K-R formula 21 methods ranged from 0.88 to 0.99 for different grades. Concurrent validity coefficients were worked out against the teacher's ratings and the annual examination marks. Vanajakshi (1970) standardized an achievement test in general science for the students of Grade VII in Andhra Pradesh. The test had a test-retest reliability coefficient of 0.90. Bhatt (1971) constructed an achievement test in science for delta class students and prepared the norms for Gujarat. The reliability coefficients found by the test-retest, split-half, K-R formula 20 and 21, and Rulon's methods ranged from 0.92 to 0.94. The SCERT, Hyderabad (1971) standardized a Science Attainment Test for Class VII of Telugu medium.

For the subjects of physics and chemistry, a few researchers like Bountra (1970), Gupta (1974), Tewari (1975), Sali (1977), Chhaya (1978), and Khandelwale (1981) standardized tests for the high school and college students. They have drawn their samples from Uttar Pradesh, Haryana, Bombay, and Maharashtra. Bountra (1970) developed achievement test in high school physics. He reported a test-retest reliability coefficient of 0.93. For the achievement test in high school chemistry, the stability coefficient was 0.95. The concurrent validity coefficients for achievement tests in physics and chemistry against the school achievement were 0.60 and 0.61, respectively. Gupta (1974) developed six different types of tests in physics and chemistry for the high school/pre-university students. The split-half reliability coefficients ranged from 0.78 to 0.95. Tewari (1975) standardized attainment tests in physics, chemistry, botany and zoology for the high school students of Uttar

Pradesh. The factorial validity of the tests was worked out. Sali (1977) standardized five Unit Tests in Physics for Standard VIII for the Maharashtra State Board of Secondary Schools. Reliability coefficients obtained through K-R formula ranged between 0.65 and 0.78. When validated against the criterion of teachers' assessment in terms of ranks, the validity coefficients ranged between 0.69 and 0.80. Chhaya (1978) standardized an achievement test in Physics for Classes VIII and X. The reliability coefficients for the tests of Classes VIII and X were 0.90 and 0.89, respectively. Khandelwale (1981) developed achievement tests in physics for Class IX in the Vidarbha region of Maharashtra. There were the first test for the first five units of the syllabus and the second test for the remaining five units. The reliability coefficients using the split-half method and K-R formula 21 were found to vary from 0.85 to 0.94. The validity coefficients against the criterion of examination marks were 0.83 for the first test and 0.85 for the second test.

Kapoor (1968) and Garg (1969) standardized achievement tests in home science for the high and higher secondary students of Uttar Pradesh, respectively. The validity coefficients against the achievement in school examinations and teachers' ratings for the test by Kapoor (1968) were 0.89 and 0.84, respectively.

Achievement tests in other science subjects have not been constructed so far. There is need to develop more and more tools on account of changing syllabi in science, non-availability of tests in some of the science subjects, and non-availability of tests at some grade levels in different parts of the country.

Miscellaneous Tests

There are some tests which have been developed for the purposes of admission, selection of subjects, and classification, etc. These were constructed by Lele and Bhagatwala (1954) for university entrance, Mascarenhas (1964) for medical fitness, Satyamurthy (1965) for subject selection, Shukla (1957) for physical education, Pillai *et al.* (1967) for handwriting, De (1965) for educational psychology, and GCPI, Allahabad (1976) for evaluation of questions. For admission to the faculties of arts and science at the university level, Lele and Bhagatwala (1954) conducted a study in Baroda. They developed the University Entrance Test. The split-half reliability coefficients of the test when administered to arts and science students were 0.91 and 0.94, respectively. The predictive validity coefficients of the test ranged from 0.54 to 0.74 for the students of science and arts faculties. With the changing nature of curricula at the university level, as well as the changing characteristics of university population, there is need for constant brushing up of the test from time to time. In order to find out the suitability of students for entrance to the medical colleges of Bombay University, Mascarenhas (1964) developed a medical fitness test. Satyamurthy (1965) developed a battery of tests for the students of multipurpose and higher secondary schools of the State of Mysore. This battery, viz., Rangachar Satyamurthy Selection Battery (RSSB), aimed at being used for guiding the students for the selection of different types of curricula. In the schools where the RSSB was used for streamlining the students, the success ratio of the schools ranged from 0.96 to 0.73. Exploring the area of physical education, Shukla (1957) developed a physical education test. He introduced the concepts of physical age and physical quotient (P.Q.). Pillai *et al.* (1967) made efforts to assess handwriting by the use of the newly standardized Kerala University Handwriting Scale. De (1965) constructed an achievement test for educational psychology covering topics like educational psychology, heredity and environment, learning, motivation, intelligence, memory, imagination, thinking, and reasoning. The reliability coefficient was 0.96. The test has suitability for teacher training institutions at various levels, school subjects, grades, regions, etc. There should be some centrally organized agencies giving guidelines for the future requirements of achievement tests.

DIAGNOSTIC TESTS

It is evident from the review of studies in the area of diagnostic tests that there have been conducted only fourteen research studies so far. Out of these, seven are Ph.D. studies and the other seven are sponsored research projects. The Bombay University, Kerala University, Kurukshetra University, Lucknow University, Agra University, Panjab University, Bureau of Psychology, Allahabad, and SIE, Gujarat, are the institutions which have conducted studies in this area. It was in 1966 that the first attempt was made by Mehta (1966), who undertook doctoral research work in the area of Diagnostic Tests. The sample in these studies varied from 300 to 3,160 students of both sexes, from Classes II to XII, and from rural as well as urban area institutions. These studies covered subject areas of Hindi, English, Gujarati, arithmetic, algebra, physics and chemistry. Diagnostic tests in the subjects of social studies, drawing, and life science are not available. From the methodological point of view, the survey method and the experimental method have been used. The types of

items included in these tests are completion, matching, alternate response, short answer, and multiple choice. The methods of K-R formula, split-half, test-retest, and parallel forms have been used for finding out the reliability of the test. Concurrent validity and content validity procedures have been used for establishing the validity of the tests.

Languages

There are five studies on diagnostic tests in languages. Sinha (1971) constructed a diagnostic test in Hindi for the Bangru-speaking students of Grade VI of Haryana. This test took into consideration spellings, pronunciation, grammar, and syntax. A diagnostic reading test in Hindi was standardized by the T.T.College, Lucknow (1971). The test was meant for the Grade III students of Uttar Pradesh. The test consisted of different parts, viz., Hindi usage, vocabulary, paired combinations, and silent reading. The split-half reliability coefficient was 0.90 and the concurrent validity coefficient was 0.69. Gomathy Ammal (1972) conducted a study to locate specific difficulties of the pupils learning Hindi in the schools of Kerala. A diagnostic test having lexical and grammatical items along with certain language skills was developed. The test had a split-half reliability coefficient of 0.84. The diagnostic test by the SIE, Gujarat (1969a) is on Gujarati spellings and use of *anuswara* for students of Grades IV and V. Sharma (1982) constructed a Diagnostic Test in English Prepositions for the students of pre-university class of Chandigarh. This was used for an experimental study of the effectiveness of density of programme and time in remedial teaching and it was found that the density and the time taken together have not shown significant results on pupils' performance.

The construction of diagnostic tests in languages at the national and regional levels needs more extensive and concentrated efforts. So far, there are only five diagnostic tests available in languages. Three diagnostic tests have been developed in Hindi for Grades III and VI of the schools of the States of Uttar Pradesh, Haryana and Kerala. There is only one diagnostic test in Gujarati for Grades IV and V of Gujarat. One diagnostic test in English preposition for the pre-university students of Chandigarh has been developed. Diagnostic tests for language skills such as listening, speaking, reading, and writing have not been developed adequately. Not even a single diagnostic test has been prepared for the English language for any school stage. It is evident from the review of the researches in this area that there is a scarcity of diagnostic tests in languages, which demands planned and systematic efforts on the part of researchers.

Mathematics

In arithmetic, Mehta (1966) constructed a diagnostic test related to the use of four fundamental rules of arithmetic. The SIE, Gujarat (1969b), also did exactly a similar work. Patel (1976) developed a Battery of Diagnostic Tests in Arithmetic for the Gujarati-medium students studying in Grades V, VI and VII in Greater Bombay. Thakore (1980) constructed Diagnostic Tests on Fractions and Decimal Fractions for students of Grade V. All these studies were followed by remedial teaching which resulted in significant improvement. Sharma (1969) constructed a diagnostic test in algebra for the students of Grade VIII of Uttar Pradesh. This was used for an experimental study of remedial teaching, which demonstrated that a high scoring student at Grade VIII could perform significantly better than a student who had failed in Grade X. Ashar (1972) standardized a diagnostic test in basic algebraic skills for the Gujarati-medium students of Grades VIII, IX and X from Greater Bombay. The median reliability coefficient was 0:90. The concurrent validity coefficients against the scores in mathematics at annual examinations of Grades VIII, IX and X were 0.94, 0.91 and 0.98, respectively. It was found that pupils committed errors due to the lack of systematic approach. The errors of conceptual type predominated those of the computational type. The trend of errors continued to a greater extent in the higher grades. To locate the difficulties and problems of students of Grades VI and VII, a diagnostic test in the Skill of Using Geometrical Instruments was developed by the SIE, Gujarat (1969). Hussen (1967), while comparing the achievement of pupils in mathematics of twelve countries reported that the achievement of pupils in this area in India is much lower than that of the pupils of other countries. To improve upon this situation deliberate and concentrated efforts are needed for improving the quality of instruction and testing services in this area. This can be improved by developing and using diagnostic and remedial programmes. Diagnostic tests in arithmetic for Classes IV to VII and in algebra for Classes VII to XI have been developed whereas, in geometry only one test in the skill of using geometrical instruments has been developed. These tests have received the attention of researchers in Gujarat, Uttar Pradesh, and Bombay. Mathematics is a subject where learning tasks are arranged in a truly sequential order; mastery of the first task is a prerequisite for the learning of the second task, and, in turn, mastery of the second task is a prerequisite

for the learning of the third task, and so on. This sequential arrangement is equally applicable from grade to grade in mathematics. The uncorrected difficulties on the sequenced tasks in Grade I are likely to be compounded by more uncorrected difficulties in Grade II, and by even more difficulties in Grade III, and so on. Hence, there is an urgent need to develop diagnostic tests in mathematics.

A general criticism about diagnostic tests is that researchers have wrongly followed the same methodology as that followed in the achievement test construction.

EXAMINATIONS

Public examinations in India play a dominating role in determining the functional content and method of instruction. They carry a good deal of prestige and act as a passport for admission to higher courses as also for entering into jobs. Research is obviously required to demonstrate that the examinations are effectively performing the functions that are required of them. An effective examination should possess four characteristics, which merit the attention of the researchers. The first characteristic of a good examination is the validity of the measures. The second characteristic is the reliability which represents consistency and accuracy of measures. Thirdly, a good examination should have a beneficial rather than harmful backwash effect upon the curriculum and teaching. Lastly, in a regional or national examination, the grading standards should be comparable.

Out of the 198 studies included in the area of Educational Evaluation and Examinations, as many as sixty-two studies have been classified in the subarea of examinations. These studies are related to various aspects of research on examination, such as achievement in annual examination, inter-examiner and intra-examiner reliability, question papers and their nature, pass percentage, external and internal assessments, correlation between theory and practical marks, mass copying, and innovations in examination. Except the Ph.D. studies by Rao (1968), Misra (1970), Deshpande (1972), Shah (1972), Tluanga (1974), Sharma (1977), Mascarenhas (1977), Singh (1978), Bhola (1978), Bhushan (1978), Singh (1978), Nath (1980), and Verma (1981), all the other studies are institutional projects conducted at Kharagpur, Gauhati, Baroda, Allahabad, Poona, New Delhi, Anand, Bombay, Annamalai, Nagpur, Saugar, Karnal, Calcutta, Pune, Madras, Hyderabad, Kerala, Jammu, and Simla. The institutional projects have been conducted by Bokil (1956-63), Shukla (1959), Lele *et al.* (1962-63), Taylor (1963), Taylor and Tluanga (1963), Taylor (1964a, 1964b), Raina (1964), Venkubai (1965), Dave and Patel (1966), Sharma (1966), Bose *et al.* (1967), Kamat (1968), Jhaveri and Patel (1968), Patel (1968), Rao (1968), Gayen *et al.* (1961-70), Harper (1962-1970), Nath (1972-74), Chatterji and Mukherjee (1973), Deo (1974, 1980), Tluanga (1974), Bhatt *et al.* (1978), Sali (1978), Mandal (1978), Sali and Umathe (1979), Koul (1979), Gunasekaran and Jayanthi (1979, 1980, 1981), Rasool *et al.* (1981, 1983), and Rajasekharan (1982).

Achievement in Examination

A significant group of studies conducted by Gayen *et al.* (1961-70) at the I.I.T., Kharagpur, have concentrated on the measurement of achievement in different subjects like English, Sanskrit, Hindi, Bengali, physics, chemistry, mathematics, geography, general science, biology, history, civics, and economics. The reference to these studies, as the title shows, could be made in the first area of the studies of this report, i.e., Achievement Tests. But these authors have not constructed and standardized achievement tests in the respective areas. They have taken the results of the school final examinations and the higher secondary examinations conducted by the West Bengal Board of Secondary Education and have applied different statistical techniques in order to evaluate the achievement in different subjects. They have taken the following points into consideration: (i) percentages of failures, passes and non-attempts for alternative question items; (ii) the grouping of items and balancing of alternatives; (iii) the difficulty values and discriminating powers of the question items; (iv) the relationship of marks in a particular item with the total marks, (v) a comparative study of the scores on one paper of a subject with another and with the total marks on the subject as a whole; (vi) content analysis of the subjects and the question papers; (vii) comparative performances of candidates in different subjects; and (viii) comparison of the internal and external assessments of students' performances. Gayen *et al.* (1961-70) found the reliability and validity of most of the present examination systems to be very low and this was so because the performance of students was assessed in a single final examination, by a large number of examiners, on a set of questions mostly of the essay type and comparatively small in number from which, again, students were given option to select alternatives. J-effect was observed having cluster of marks at pass mark and other critical zones, followed by big gaps just below these.

Examiner Reliability

That traditional essay-type examinations are unreliable has been proved again and again. Researchers like Harper (1962), Jhaveri and Patel (1968), Misra (1970), Taylor (1963, 1964a, 1964b), Nath (1972), Deo (1974, 1980), Tluanga (1974), Bhushan (1978), Bhatt *et al.* (1978), Gunasekaran and Jayanthi (1979, 1980), and Verma (1981) conducted studies related to inter-examiner and intra-examiner reliability. Three studies by Harper included in this review are (i) research in examinations, (ii) chance in the traditional examinations, and (iii) objective and traditional examinations. In the study, Chance in the Traditional Examinations, it is pointed out that chance plays a role in determining the achievement level of a pupil. In another study by Harper (1962), it is found that objective examinations are more reliable and valid than traditional examinations. Harper (1970) conducted his ninety marking ten experiment. Ten history answer-books in which the candidates had answered the same five (out of ten) questions were selected. It was found that examiner's marks of these ten answer-books ranged from 2 to 38 out of 50. For one examiner the pass percentage was 80, while for the four it was only 10. The average of all the 90 examiners for the ten answer-books ranged from 8.8 to 27. It was found that inter-examiner reliability was very low. The low inter-scorer reliability was attributed to the essay type of examinations, types of questions, inadequate instructions and training of examiners. Misra (1970) found inter-examiner reliability coefficient of 0.70 for essay. Taylor (1963) made an extreme comment that an examiner's marks had neither the sanctity nor the precision which was usually attached to them. Different examiners showed a large variation in the mean and standard deviation of their evaluation. Taylor (1964a), in his study, An Examination of Examiners, emphasized the same point of low inter-examiner reliability. He, therefore, suggested that marks should be scaled to the same mean and standard deviation before they are combined. He proposed and then experimented with the technique of random assignment of roll numbers in the diagonal order with a rectangular grid. He found that by this technique more candidates could pass the examination. He said that it was much more likely than before that these candidates were the ones who ought to have passed. Bhushan (1978) studied various methods of examination in history at the lower and higher stages. The evaluation of ten examiners each at lower and higher stages was utilized in the assessment procedures of essay format, short answer format, open book examination, viva-voce and objective type tests. It was found that there was divergence in marking standards at both the levels in essay type, short answer, and open book type of examinations. At the higher stage, short answer type tests worked well, whereas at the lower stage essay type of examination worked well. Bhatt *et al.* (1978) conducted a comparative study of the grade table method (GTM) and direct grading method (DGM) and found that three out of the four examiners were consistent within themselves in respect of grades under GTM, whereas each of the four examiners showed consistency within himself in grading the scripts under DGM. They found that examiners were consistent between themselves when the scripts were ranked in order of merit under GTM as well as under DGM. Gunasekaran and Jayanthi (1979) conducted a study to fix the permissible differences in grades awarded by two examiners under the direct grading method. They stated that the permissible differences between grades assigned by two examiners were fixed as one grade. Verma (1981) studied the innovation of marking versus grading in examination. For this, fifty answer scripts of B.A. (History) were got evaluated on a five-point scale and seven-point scale from five examiners. It was reported that there was no significant difference between average assessment on 5 and 7 points letter grade scale, but better discrimination was possible through the latter. In this context two studies of Deo (1974) and Tluanga (1974) are interesting to note. Deo (1974), in her study, Effects of Revaluation on the Results of Candidates Appearing at the University Examinations, found that out of 484 candidates the results of as many as 240 students changed significantly. Deo (1980) included all the universities except agriculture and technological universities in the sample and reported that: (i) more failed candidates applied for revaluation than already passed candidates, (ii) the proportion of marks-increased cases in the failed group was larger than marks-increased cases in already passed group, and (iii) the correlations between the original and revalued marks were very high, establishing a high reliability of assessment of essay-type questions. Gunasekaran and Jayanthi (1980) conducted a study of the revaluation cases. It was found that about 25 per cent of the applicants benefited as a result of revaluation. On the aspect of multiple examiners, Tluanga (1974) suggested that there was no justification for referring the script to a third examiner on the ground that the first two examiners gave divergent marks. The intra-examiner reliability, too, is a matter of concern in essay type of examination systems. Jhaveri and Patel (1968) found that intra-examiner reliability could be increased

through well-defined essays rather than traditional essays. Nath (1972) also found that scaling reduced the dispersion of marks distribution, thus bringing down the variations between the examiners. In another very interesting study, Supplementary Examinations, Taylor (1964b) reported that a supplementary examination always increased the proportion of bad candidates in the pass list by a factor which was likely to be between 1.5 and 2.0. The GCPI (1981) investigated the appropriateness of the supplementary examinations and reported that they should be abolished.

Question Papers

Studies by Lele *et al.* (1962b, 1963a), Chauhan (1967), Rao (1968), Malhotra (1972), the GCPI, Allahabad (1971), Gnanapragasam (1975), the SCERT, Hyderabad (1977), Mascarenhas (1977), and Bhola (1978) have dealt with question papers and their nature. Chauhan (1967) conducted a broad study related to the university examination system. He found that of the students who failed a large proportion failed not because of not knowing the subject matter but because of some external factors like defective question papers, carelessness or examiners in evaluating answer-books and evaluation by incapable examiners. In a similar context, Malhotra (1972), while studying the effectiveness of question papers of matriculation examinations, found that the question papers had many defects with respect to difficulty level, coverage and weightage. As regards the type of questions used in our examinations, Lele *et al.* (1962b) analysed question papers and found that only one-fifth of the total questions were good, whereas the rest were poor discriminators. Rao (1968) investigated the system of tests and examinations at Standard XI. He found that in many cases question papers were not properly balanced as far as the difficulty value of the items was concerned. Question papers even failed to discriminate pupils of high and low ability. In a similar context, Bhola (1978) also found that question papers were not well balanced in respect of attributes of discriminating power, difficulty value, reliability, and validity of the question items. He further found that alternative question papers set for the morning and the evening sessions were not analogous and of same standard. Furthermore, Lele *et al.* (1963a) found that in essay-type examinations teachers and students did not agree as regards the difficulty level of the questions and that the selection of questions on the part of students varied with the nature, clarity and difficulty level of questions. It is interesting to find out the functioning of items in a question paper. Related to this aspect Sali (1978) conducted a study to analyse student's achievement in each constituent question of mathematics in the secondary school leaving examination. It was found that out of 45 subitems in nine questions, only three subitems were responded to by more than 80 per cent candidates, 14 subitems were responded to by only 20 per cent candidates and the remaining 28 subitems were responded to by 80 per cent candidates. The GCPI (1971) conducted an experiment in reforming the examination system by designing question papers and allowing students to use books, notes, etc., in the examination hall. Accordingly, suitably reformed question papers for Grade VIII were prepared for the subjects of English, Hindi, mathematics and general science. The questions based upon the direct or indirect reproduction of facts, learnt through memory, had no place in the reformed question papers. Such questions were set as required creative thinking, proper understanding and comprehension of facts. The study revealed that open book examination did not unduly help the students when question papers were set in a reformed way. Mascarenhas (1977) analysed the question papers of past ten years in higher level English and in geography with the purpose of seeing their strengths and weaknesses. It revealed that knowledge aspects dominated only a few items testing application or interest in both the subjects. The SCERT, Hyderabad (1977), while studying the pattern of question papers of Grade VII examination revealed that nearly 82 per cent of the examiners felt that the essay-type questions allowed greater discretion for them while 21.2 per cent said so in the case of short answer questions and 17.6 per cent said so in the case of objective-type questions. Bose *et al.* (1967) reported that a question paper should include double the number of questions to be answered. These questions should be uniformly distributed throughout the syllabus. Irrespective of the number of questions in a question paper, the most difficult task faced by the paper setter was related to the framing of good questions. In this respect, Hill (1964) stated that it was beyond the competence of even the most skilled paper setter to make each question in every question paper function exactly as he wished it to do.

Question Banks

Singh (1978), while validating Bloom's taxonomy of educational objectives by analysing the product and process-oriented approaches, found that (i) the hypothesis of increasing complexity of the mental processes involved in the objectives ranging from knowledge to

evaluation, was supported, the objective, synthesis, seemed to be misplaced while the objective, analysis appeared to be misleading. Based on Bloom's taxonomy of educational objectives a question bank consisting of 2,983 objective-type questions/items was developed by Gnanapragasam (1975). Natrajan through the Association of Indian Universities has developed question banks at the undergraduate level in several subjects, namely, mathematics, chemistry, botany, geography, economics, physiology, political science, the English language and literature, automobile engineering, physics, zoology, history, psychology, commerce, sociology, foods and nutrition, pharmacology, and law of contracts. Also, a question bank was developed in mathematics at the postgraduate level. At the undergraduate level, question banks are under preparation in other subjects, namely, anatomy (medicines), constitutional law (law), strength of materials (engineering), electrical power system (engineering), and library classification (library science). At the postgraduate level, a question bank in English literature is also under preparation. In education, Hooda (1976) prepared about 300 multiple choice items. Later on, the NCERT and the Examination Reform Unit, Madras University, made attempts to develop question banks in education.

Pass Percentage

Difficulties regarding assessment through the present system of examinations are manifold. Researches should be conducted to settle such important issues as awarding of grace marks, fixing the pass-fail cut-off score, and so on. Dave and Patel (1966) found that there was a substantial variation among the pass percentages in school examinations in a given year. They found a substantial increase in variation in the pass percentages over a five-year period from 1960 to 1964. Sharma (1966) analysed the public examination results of Uttar Pradesh Board and found that, from year to year, the average performance continuously increased in Hindi, mathematics, and science but in English a reverse trend was observed from 1961 to 1965. The SIERT, Karnal (1970), found that there was a substantial increase in variation in the pass percentages over a five-year period from 1964 to 1968 of middle standard examination results. Rajasekharan (1982) compared the results of five south Indian universities, namely, Kerala, Calicut, Madras, Mysore, and Tirupati, and found that there was significant variation in the percentage of passes over subjects and years. An important reason identified for the variation was the nature and difficulty level of the questions. Shukla (1959) found that the levels of marks obtained by examinees of different subjects at the M.A. Examination of the University of Delhi in the years 1950 to 1957 had different meanings. Nath (1974), in a study of the results of some selected colleges of Gauhati University, found that the subjects of economics, English and geography contributed significantly to label a college as below average in attainment.

Internal-External Assessment

A few studies aiming to find out the relationship between external and internal assessments and also their reliability were conducted by Raina (1964), Kamat (1968), Deshpande (1972), Shah (1972), Gunasekaran and Jayanthi (1980), Nath (1980), and Rasool *et al.* (1981, 1983). Shah (1972) surveyed the pattern of internal-external assessments in arts, commerce and science colleges. He found that there was a tendency to decrease weightage of internal assessment in arts, science and commerce colleges, while in colleges of education the tendency was in favour of increasing internal assessment. He found significant correlation between internal and external assessments. Deshpande (1972) also found a positive correlation between internal and external assessments, but it varied from school to school. Kamat (1968) also conducted a study related to internal and external assessments. He found that correlation between internal and external assessments was not very high; it was less in arts than in science subjects; it was less in newer colleges and in mofussil colleges than their counterparts in Poona City. Rao (1968) also found that correlation coefficients between the external and the internal assessments were greater in language subjects than in other subjects. Gunasekaran and Jayanthi (1980) found significant correlation between the marks of continuous internal assessment and the university examination. Nath (1980) found that the internal assessment marks possessed some predictive value for the external assessment marks but there was a tendency towards overmarking in internal assessment. Rasool *et al.* (1981), while studying internal and external awards at the postgraduate level in Jammu University, found that the majority of the teachers preferred lower range in the marks of the internal assessment than that of the external assessment. Most of the coefficients of correlations were positive between internal and external examiners. Again, Rasool *et al.* (1983) studied the trend of marking in internal and external assessments. They found that the trend of marking was highly liberal under internal assessment. The percentage of first divisioners in overall

result was higher and the percentage of third divisioners and failures got reduced due to the addition of internal assessment. Venkubai (1965) found that nearly 30 per cent of the schools of Hyderabad and Secunderabad resorted to inflation of marks in the internal assessment of 25 per cent weightage in mathematics and science. But internal assessment was not found to be responsible for higher percentage at the public examination. In a seminar on Examinations in Higher Education, Bennur (1971), Misra (1971), and Tare (1971) presented papers related to internal assessment. Bennur (1971) reported correlation coefficients between internal and external assessments for two years as 0.41 and 0.56 for B.Ed. examinations, 0.04 (not significant) for the second year engineering mathematics, and –0.38 (not significant) for the M.A./M.Sc. mathematics students. The only study related to the attitude of teachers towards internal assessment was conducted by Reddy (1979). He found that teachers of the college where internal assessment had not yet started did not differ in their attitude towards the internal assessment system from those of the college where the system was in practice.

Over many years, researches related to the issue of internal and external assessments have been conducted at the school level and the college level, in different faculties, in different geographical areas, and with different types of examination systems. The analysis of results has been carried out by employing mean differences, pass-fail differences, and correlations. Most of the results have indicated that there was liberal marking in internal assessment and it had varying relationship with external assessment. The educational researcher should suggest workable ways and means to solve the issue of internal assessment on the basis of planned research.

Theory and Practical Examinations

Very few studies have been conducted to find out the relationship between the marks obtained in theory and practical examinations in different subjects at different levels. Sali and Umathe (1979) studied the relationship between the marks obtained by students in theory and practical examinations in science at the S.S.C. Examination of Maharashtra. It was found that the mean obtained score was 38.51 out of a maximum of 120 theory marks. Further, they found that the obtained mean score was 21.61 out of a maximum of 30 practical marks. The coefficient of correlation between the marks in theory and practicals was found to be 0.34. Only 6 per cent of students secured more than 60 per cent of marks in theory, whereas only 17.5 per cent of students secured less than 60 per cent of marks in practicals. This showed that there was a wide range of marking in theory and practical aspects of examinations. Hooda (1976) studied the relationship of attainments in theory papers with the student teaching in secondary training colleges. He found that the coefficients of correlation of the student teaching scores with the scores in the theory papers ranged from 0.03 to 0.33. These studies revealed that marking in theory and practical aspects of examinations had a wide disparity. More research studies in the future should be conducted with a view to settling the issue of marking pattern in theory and practical examinations.

Mass Copying

In addition to the above issues of internal and external assessments, relationship between theory and practical marks, reliability of examination, question papers and their nature, etc., one more recent issue, namely, mass copying, has emerged on the Indian scene. It has become important to study the genesis/sources of mass copying. In this context, Mandal (1978) conducted a study to find out the factors associated with the mass copying behaviour of students in the final secondary examinations with particular reference to their home and school conditions. The study involved ninety-two non-copying students from non-mass copying schools and ninety-two copying students from mass copying schools. The study revealed that (i) the mass copying students had relatively poor socio-economic status, low parental education, and inadequate facilities at home, (ii) the non-mass copying institutions differed from the mass copying institutions in respect of faculty subsystem, administrative subsystem and student subsystem.

Innovations

Some experts in the country have given a few suggestions to improve the traditional examinations. Taylor (1963) suggested the scaling of marks. Bose (1965) suggested the screening of examiners. Harper suggested the appointment of independent examiners for each question. Taylor and Tluanga (1963) suggested the use of standard error or marking instead of arbitrary principle of grace marks, and others suggested the vertical system of promotion of individual pupil, stage by stage and subject by subject. These issues have to be validated by research studies in the future. Koul (1979) studied the acceptability and feasibility of thirteen innovations in the examination system at the university stage. He found that a few innovations, namely, introducing the grade

system of ranking, supplementing essay-type examination with objective-type examination, semester system, internal assessment supplementing the external examinations, and defining the scope of questions in simple and clear language, were acceptable to students and teachers. Many more innovations for improving examinations have been suggested. The introduction of these innovations requires careful examination and preparation. Researchers should pay adequate attention to this area.

FACTORS AFFECTING ACHIEVEMENT*

In this area nineteen studies have been included. These studies primarily focus on research in examination and evaluation. But they also cover some of the correlates of achievement and factorial structure of achievement domain.

Correlates of Achievement

There have been studies investigating the correlates of achievement, such as, the medium of instruction, number of languages learnt, age, sex, caste, and locality of students, nature of school, size of school, type of curriculum offered, student teacher ratio, nature of question paper, nature of examinations — theory or practical, and day scholars or hostel residents. All these variables can be classified into four types, viz., tester, testee, testing situation and the test.

Regarding the tester variables, studies have been reported earlier in the section on examinations where it has been brought out that the examiner is a very important factor influencing examination marks. Misra (1970) said that out of ninety examiners scoring the same ten answer sheets, one of the examiners passed only 10 per cent of the students, whereas another examiner passed 90 per cent of the students. Patel (1967) conducted a study on the results of practical examinations and found that inter-examiner reliability was one of the serious issues.

The testee variables like the age, sex, caste, and locality of students in relation to achievement were studied by investigators like Bokil (1956a), Dave (1958), Sharma (1967), Sharma (1976), Sharma (1977) and Rawat (1978) in Maharashtra, Gujarat, Rajasthan and Delhi. Bokil (1956a) found that the pass percentage of the girls in the S.S.C. Examination was more than that of the boys. Dave (1958) found that girls scored a little lower in arithmetical abilities than boys up to Grade VI, but more in Grade VIII. In the case of vocabulary test, girls scored more than boys at the age of eight and nine years. Sharma (1976) reported that girls were found to be significantly superior to boys both in general science and mathematics. Contrary to the above findings, Sharma (1977) found that there was no significant difference in the achievement of boys and girls in arithmetic and in the subject of copying of designs, but in reading test girls achieved significantly higher scores than boys. Bokil (1956a) found that in the case of the S.S.C.E. candidates, the pass percentage decreased with the increase in age to the extent that at sixteen, the pass percentage was 77 and at twenty, the percentage was 27. Sharma (1967) found in the context of caste of students that with higher levels of caste there was better achievement. Bokil (1956a) found that the number of languages learnt affected achievement. Rawat (1978) found that there was positive correlation between the reading ability and achievement of pupils in different school subjects and the general academic achievement. He also found that the rural children showed poorer reading ability than their urban counterparts.

As far as the nature of test instruments and its influence upon examination marks is concerned, a large number of studies are reported earlier under the heading Examination. Related to testing situation very few studies have been conducted. The GCPI (1981) studied the factors responsible for good and poor percentage of examination results. It was found that proper facility of library, reading room, laboratory, playground, etc., helped in increasing the percentage of examination results. The other important factors which helped in increasing the percentage of examination results were the teaching experience of the staff and principal, the methods of teaching, regular correction of homework, regular evaluation, proper guidance and encouragement to students, healthy relationship between the principal and the staff, cooperation between teachers and parents and good management. The various factors responsible for the poor examination results were lack of dedicated teachers, lack of interest in studies among students, lack of proper correction of homework, lack of material resources in the school and the passive attitude of parents towards the education of their wards. Investigators like Bokil (1956b) and Sharma (1967) found that the variable medium of instruction influenced achievement significantly. But the SIE, Maharashtra (1971) found no significant difference in the achievement of students with English and non-English as their media of instruction.

*More studies related to the area of correlates of achievement have been reported by Anand and Padma in Chapter 13 of this publication — *Editor*

Bokil (1956a) found that the pass percentage of students appearing for the S.S.C. Examinations held in March and October differed significantly. Bokil (1959b) also found that pass percentage was the highest in schools of medium size. Nath (1974) found that the achievement of pre-university students was related to student-teacher ratio. Sharma (1967) found that the achievement of high school day scholars was better than that of night school scholars. Shukla and Tuttoo (1959) compared the achievement of children of different types of schools (basic and non-basic). They did not find a clear-cut superiority of one type of school over the other. Sharma (1967), Bokil (1956b), Rao and Arunajatai (1971) found that the nature of curriculum and subjects offered by students influenced their achievement scores. Malhotra (1972) found that pass percentage differed from question to question in different papers at the matriculation examination of the Panjab University. The above studies show that testing conditions and the nature of test influence the achievement of students.

Factorial Studies in Achievement

The nature of factors expected within the domain of achievement is dependent upon the variables and tests, sample size and its nature, testing conditions and methodology of factoring. Factor analysis as a technique has been used by Chickermane (1943), Gokhale (1954), Mehrotra (1954), Lele *et al.* (1964a), Pendharkar (1965), Dash (1967), Sinha (1967), Kamat (1968), Deshpande (1972), Chatterji and Mukerjee (1973), Gupta (1974), Tewari (1975), Singh (1978), Singh P.S. (1978), De (1979), and Chauhan (1982). Mehrotra (1954), using the marks in the subjects of both high school and intermediate examinations found five factors out of which there was one factor named as General Scholastic Factor, the other four factors had high loadings on languages, social studies, mathematics and science, fine arts and crafts. Dash (1967), while factorizing his Achievement Test Battery for Oriya, mathematics, social studies, and general science found four centroid factors, viz., General Scholastic Ability, Verbal Application, Mathematical Ability, and Memory. Lele *et al.* (1964a) factor analysed the correlation matrices of different subjects of S.S.C. and P.Sc. examinations taken for three consecutive years. On examining the centroid factor matrix, it was found that the first two factors, which shared most of the common variances, had stability in different samples of three consecutive years. These two factors were named as Scholastic Aptitude—Verbal and Numerical, and Scientific Factors. While analysing the correlation matrix of different variables related to achievement in English, Sinha (1967) identified three factors, viz., Automatization, Comprehension-Vocabulary, Pronunciation and a fourth factor having loadings on grammatical aspects of English achievement. As regards the study in arithmetic, Chickermane (1943) identified Reasoning (G), Number (N) and Verbal (V) factors. For geometry tests, Gokhale (1954) found General, Number and Spatial factors. Kamat (1968), while analysing the internal and external assessments marks, found three factors underlying the domain of achievement. Gupta (1974) factor analysed the attainment of pre-university students in the subjects of physical sciences and mathematics. He found that five factors could explain the domain of physical sciences and mathematics. These were: (i) general mathematical ability factor, (ii) general factor of scientific ability, (iii) symbol and number factor, (iv) deductive reasoning factor, and (v) space factor. Tewari (1975) also factor analysed the attainment in the areas of science and mathematics. He could extract five factors, namely, (i) theoretical factor, (ii) mathematical factor, (iii) biological science factor, (iv) chemistry-botany factor, and (v) general science factor. Chatterji and Mukerjee (1973) factor analysed the correlation matrix of annual examination marks of different subjects of Class VIII. On examining the centroid factor matrix, it was found that most of the school marks on all the seven subjects measured one general factor which was identified as Scholastic Achievement. Singh (1978) factor analysed the marks in preparatory science examination in theory, tutorial and practical. In theory, as well as in tutorial examinations three common factors (a) general scholastic, (b) scientific verbal, and (c) scientific numerical emerged, whereas in practical examination only one common factor of general ability was found. When the marks of all the three assessments were added, one more factor, viz., problem-solving, was found. Singh (1978) found that factor analysis did not support the cumulative nature of Bloom's taxonomy as only one general factor was traceable. De (1979) factor analysed the scholastic achievement in geography for the secondary students of West Bengal and found three principal factors, viz., Geographical Reasoning, Geographical Aptitude and Geographical Application. Chauhan (1982) factor analysed the correlation matrix of eight selected structures of intellect abilities as supposed to be involved in algebra and eight tests in algebra having different steps necessary for algebraic problem-solving. He identified two factors, general algebraic reasoning and pattern visualization.

PREDICTION-ADMISSION-PROMOTION STUDIES

How effective are the predictions? There can be three approaches to handle this issue viz., (i) simple correlation between the predictors and the criterion, (ii) multiple correlation, and (iii) factor analytical approach. A limitation of such studies is that the investigators take the sample of the candidates who have been admitted to a concerned course. The investigator can know the success-failure ratio of those who have been admitted on the basis of some predictors, but they cannot find out success-failure ratio of those who have not been admitted. It is in the context of this limitation that the findings of prediction-admission-promotion studies should be understood.

Different types of predictors have been used in the research studies included in this trend report. Mathur (1971) found that Jalota's General Mental Ability Test was a good predictor for general science ($r = 0.80$). Nafde's Non-verbal Intelligence Test was found to be an effective predictor for mathematics. He also used other tests like Saxena's Personality Inventory, the DAT, and the Numerical Aptitude Test. Satyamurthy (1965) used a battery of psychological tests for selection to multipurpose and higher secondary schools and found that success ratios in respect of diversified courses ranged from 0.93 to 0.80. Chakrabarti (1974) developed a test battery consisting of scientific knowledge, technical knowledge, commercial knowledge, and mathematical knowledge for predicting probable success of students in the academic streams of humanities, science, commerce and technical. Singh (1975) developed a test battery consisting of manual dexterity, cause-effect relationships, spatial perception, reasoning, observation, numerical ability and memory for predicting success in achievement in science at the intermediate level. The multiple regression coefficient was 0.48. Ramachandran *et al.* (1971) studied the prediction of scholastic achievement in two major streams through multiple discriminant function based on the interest scores. They found that students could be classified with considerable precision into science and humanities streams on the basis of their interest pattern. Mitra *et al.* (1978) found that interest in science could be employed as a differentiating factor in the identification of more capable science students.

Bokil (1958b), Buch (1963) and Sengupta (1963) used the final examination marks for predicting achievement in the next final examination. Studying the prognostic value of higher secondary examination marks of Delhi, Dutt (1954) found that the subject choice at the degree level by those who secured less than 45 per cent marks should be made with caution. Higher secondary examination marks in languages and mathematics were useful for the selection of courses and prediction purposes. Dharm Bir and Srivastava (1965) tried to find out the predictive value of matriculation examination for the intermediate and degree examinations. They found that subjects like Sanskrit, mathematics, arithmetic, and household accounts had higher prediction coefficients than subjects like geography and English. Bhatt and Indiresan (1981) tried to find out the predictive value of matriculation examination for the achievement in polytechnic and found that the high school scores were not a good predictor of success in polytechnics and hence, the high school scores as the only basis for admission need to be reconsidered. Buch (1963) found that S.S.C. Examination composite scores of physics, chemistry, and mathematics were better predictors than the individual set of scores while predicting success in the next higher examinations. Palsane (1965) studied the predictive value of the S.S.C. Examination for the degree examination in the faculties of engineering, arts, commerce and science. He found that the language group marks had the highest predictive potentiality in the faculties of arts and engineering. He also found that in the faculty of commerce the marks in social studies at the S.S.C. Examination appeared to be good predictors; in the faculty of engineering, the marks in mathematics as well as science at the S.S.C. Examination had high predictive value; and in the faculty of science, the marks in science subjects at the S.S.C. Examination appeared to be a poor predictor. Hiriyanniah (1963), while finding out the prognostic value of pre-university examinations, found that the science students were relatively consistent in their level of attainment at later stages in the university than the arts and commerce students. Hence, the predictive value of the pre-university examination in science was better than that in arts and commerce groups. Nath (1973) used M.A. (previous) examination marks for predicting success in the M.A. final examination for the subjects of English, economics, and political science. The correlation coefficients were high and ranged from 0.76 to 0.83. Lele *et al.* (1962a), while conducting a study to ascertain the relative efficiency of scaled and unscaled marks for predicting future success, found that the unscaled marks turned out to be as good predictors as the scaled marks.

It may be noted that only one study by Lele and Bhagatwala (1954) has developed a standardized entrance test. The test had a maximum predictive validity coefficient of 0.74 for arts students. Narayana (1979)

conducted a study of the performance levels of candidates appearing at the Andhra Pradesh Public Service Commission Examinations. The selection of candidates with first division predominated over the candidates with second division and the selection of the candidates with second division predominated over the candidates with third in all types of services. If the nature of students and the nature of predictor and criterion variables remain stable, prediction studies will be more useful, but, unfortunately, it is not the case. Besides this, if the correlation coefficients between predictors and criterion variables are negligible, and/or the academic success is specific and situational in nature, then the prediction studies are not very useful and safe. With this background in mind, future prediction studies may be designed and undertaken.

FAILURES

This review includes twelve studies which are directly related to the phenomenon of failures. The sample size in these studies varies from 1,000 to 28,449 candidates. The samples include boys and girls, from rural as well as urban areas, belonging to Classes VI to XI. The studies cover government, aided, private, local body and non-aided institutions. Out of the twelve studies one is a doctoral study and eleven are institutional studies related to the area of failure in examinations. Bokil (1956c, 1958a, 1963) and Gadgil (1978, 1979a, 1979b) conducted studies by drawing samples from the S.S.C. Examination Board of Maharashtra. The five institutional studies of the DEPSE (1964), the NCERT (1964, 1965), the GCPI, Allahabad (1964), and the Directorate of Higher Education, Hyderabad (1966) are related to the area of failure in examinations. Kumar (1982) conducted a study to investigate failures in Kendriya Vidyalayas.

Pass-failure in an examination depends upon the interplay of a large number of variables related to pupils, teachers, subjects, schools, the examination system, and parental status. Studies related to failures have been conducted at the school level as well as at the university level. Bokil (1956c) found that the school size was not so important a factor which affected the percentage of failures, but the size of school influenced the number of candidates scoring less than 20 per cent of the total marks. As for the locality of school — rural or urban — Bokil (1956c) found that rural schools had more variation in failures. While conducting studies in relation to failures in English, Bokil (1963) found that one-third of the failures were by a margin of less than 10 per cent marks; overall average score in English was thirty-one, which was less than the pass marks of thirty-five. Bokil (1958a), while calculating average percentage of failures in English, found that there were significant differences from group to group and from year to year. The DEPSE (1964) found that teachers' qualifications and background, teaching methods, working conditions and the location of schools, transfers, building, equipment, clerical work done by teachers, pupils' previous attainment, pupils' attendance, media of instruction, examinations, etc., were the factors which were related to pass-fail percentage of schools. It was also found that the largest failures at the S.S.C. level were in the subjects of English and mathematics. There was a significant difference between the pass percentage of private and regular candidates. The GCPI (1964) also found favourable results for regular candidates. It was found that girls failed less than boys. The performance in English, mathematics, science, and civics was responsible for the higher incidence of failure. The study found that many social, economic, and educational factors were responsible for failure in schools. The NCERT (1965) conducted a sample study of failures in the boards of secondary examinations taking the sample from the boards of Bihar, Delhi, Gujarat, Kerala, Maharashtra, Mysore, Rajasthan and Uttar Pradesh. It was found that the compartmental examinations conducted by certain boards did not seem to make much difference to the large-scale failures. It was further found that the majority of failures were not due to failure in English alone. The Directorate of Higher Education, Hyderabad (1966) found that the high percentage of failures in the examination at the secondary stage was mainly due to inadequacy of library and laboratory facilities and untrained teachers. While conducting studies in relation to failures in English, mathematics, and social science, Gadgil (1978, 1979a, 1979b) found that a large number of failures were mainly due to inadequacy of mastery over the subject by the teachers.

There are significant gaps in research efforts in this area. Indepth studies are needed to have better knowledge about this aspect of education.

Summing up

This section includes some outlines of the gaps, overlaps and suggestions for research priorities in the area of Educational Evaluation and Examinations. These outlines have mostly emerged from the perusal of literature and foresight. The researches so far conducted in the field of Educational Evaluation and Examinations can be characterized as inadequate, not well integrated and,

therefore, at moments, not goal-directed.

The researches appear to be inadequate when the importance of the field and the existence of gaps are considered. The gaps are conspicuous by the absence of studies related to achievement tests in regional languages, some science subjects, new curricula of mathematics and science, diagnostic tests in almost all the school subjects, failures, etc. Research studies aiming at achievement test construction at the primary stage and the higher education level are almost negligible. There have been instances when studies are duplicated with almost the same content, methodology, region, and even time. Duplication is also there when most of the prediction studies, covered in the review show a trend of using the same type of predictors and the same type of criteria. The second characteristic of Indian researches in the field of educational evaluation has been labelled above as not well integrated. This is explicit by the arguments given to show the gaps and overlaps. To argue further, such a trend is explicit because no individual or institution has undertaken well-knit studies, over a longer period, covering a large number of school subjects and sample throughout the country. Only one or two studies are exceptions to this argument. Thirdly, it is viewed that the studies are not goal-directed because the construction of achievement tests and evaluation procedures has been treated as an end in itself. Probably what is required is that these tools should be used to provide feedback to students, teachers, administrators and planners for establishing the objectives, framing the curricula, promotion, admission and classification of students and for research purposes. This may be difficult for individual researchers and autonomous institutions. But it is felt that there is need for such a programme. Who can do it? How can it be done? These are the questions for the future researchers and administrators to answer.

To add another characteristic to the Indian researches in the area of educational evaluation and examinations, it may be stated that most of our studies are conducted by following the traditional steps of tool construction within the framework of norm referenced testing. New concepts like criterion referenced testing and mastery learning have to be incorporated in tool construction wherever necessary. The review of the Indian studies indicates that the samples used are small and at moments not properly drawn. The variations about the sample size and the methodology of sampling are very wide. In the present studies, two item indices, namely, item difficulty and item discrimination, have been widely used. It is also observed that researches have not explicitly stated the nature of the sample for item analysis. It was observed that in most of the studies related to essay-type tests, the content reliability has not been computed for establishing the examiner reliability. Regarding the method of establishing concurrent validity of achievement tests included in this review, the criteria used for validating achievement tests suffer from defects mentioned earlier. The present researches do not indicate the factors influencing the validity of their tools. These deficiencies create doubts about the scope for generalizability of the findings. It is, therefore, suggested that studies in achievement testing should be designed so as to maximize the generalizability of the findings. It is suggested that research studies should be planned to fill up the gaps in the area of achievement tests. General scholastic achievement tests have not been developed for the subjects of social studies, economics, physics, chemistry, civics, regional languages, etc. Efforts should also be made to develop general scholastic tests related to basic skills and general educational development, so that they may be used for a variety of purposes. Achievement tests in regional languages have been developed in Gujarati, Oriya, Kannada and Marathi. Achievement tests for subjects like economics, sociology, etc., have not been standardized at all. It is also observed that with the change of curricula in mathematics and science, the available achievement tests will be outdated. There is need to develop more and more tools on account of the changing syllabi in school subjects. On the basis of the review of studies in this area, it can be safely remarked that there is a lack of planning and coordination in the work done. Systematic and continuous work over a period of time should result in workable tools for different school subjects, grades, regions, etc. Some centrally organized agencies giving guidelines for future requirements of achievement tests should be established.

It is relevant to mention here that there is not only a scarcity of diagnostic tests in the country, but also the need to improve the diagnostic tests. So far, there are only five studies employing diagnostic tests in languages, such as Hindi, Gujarati, and English. Diagnostic tests for language skills, such as listening, speaking, reading, and writing, have not been adequately developed. Not even a single diagnostic test has been prepared for the English language for different school stages. Diagnostic tests in arithmetic for Classes IV to VII, in algebra for Classes VII to XI, and in geometry in the skill of using geometrical instruments, have been developed. The gaps are conspicuous by the absence of studies related to diagnostic tests in science subjects, social sciences, and

some languages. Research studies aiming at diagnostic test construction at the primary stage and the higher education level are almost missing. It is suggested that diagnostic tests be constructed by following appropriate methodology. It is also suggested that planned and systematic efforts for the development of diagnostic tests should be made by the NCERT and state-level organizations.

There are sixty-two studies related to various aspects of research on examinations, such as achievement in annual examination, inter-examiner and intra-examiner reliability, question papers and their nature, pass percentage, external and internal assessments, correlation between theory and practical marks, mass copying, and innovations in examinations. Research studies in examination should be undertaken to fix the proportion of objective, short answer, and long answer questions in every examination paper. While it was suggested in the Mahabaleshwar seminar of 1981 that the proportion of objective, short answer, and long answer questions can be 25 per cent, 35 per cent and 40 per cent, the actual proportion can be decided by every examining agency on the basis of research. Every university and Board of School Education should conduct studies related to the choice of questions in question papers, internal opinions in essay questions, etc. Studies related to marking in theory and practical have revealed that there has been a wide disparity. More research studies should be conducted in the future with a view to settling the issue of marking pattern in theory and practical examinations. Studies should also be undertaken to develop strategies for the implementation of the innovative programmes like open book examinations, internal assessment, grading system, scaling of marks, question banks, national examination, question-wise marking, central paper setting, etc. Research studies should also be conducted related to the development of the administrative infrastructure and effective management of examinations. National Testing Service, with a research cell, should be established for conducting research, as well as to implement innovative programmes in the area of Educational Evaluation and Examinations.

In the area of factors affecting achievement, studies related to the correlates of achievement and factorial structure of achievement domain have been reviewed. Research studies related to the correlates of achievement at the primary stage and the higher education level are almost negligible. The studies reviewed in this area show that testing conditions and the nature of test influence the achievement of students. The nature of the factors expected within the domain of achievement is dependent upon the variables and tests, sample size and its nature, testing conditions, and methodology of factoring. The correlational and factorial studies have tried to describe the correlates of achievement and factorial structure of achievement. Both these efforts have been inadequate and sporadic. It is necessary to conceive different paradigms for different types of educational achievements. It must be mentioned here that very comprehensive paradigms cannot be easily verified. Hence, the sectorial approach to different domains of educational achievement needs to be studied through factor analytical studies. Subsequent to these sectorial studies, efforts should be made to integrate the specific factor structures, by linking inter-domain, inter-stage, inter-region, and inter-examination system structures. Such efforts may lead to more generalized hierarchical structure or morphological structures on the lines of intelligence. Moreover, the factor structure of achievement under convergent and divergent test contents and format is yet to be explored.

While screening the existing prediction-admission-promotion studies, four points emerge: (i) most of the prediction studies are retrospective rather than prospective; (ii) while finding success-failure ratios, it is only the admitted selected samples which have been used and their counterparts, i.e. the rejected candidates who wanted to join particular courses, have been ignored; (iii) examination marks which probably themselves lack the characteristics of reliability and validity have mostly been used for prediction; and (iv) psychological profiles consisting of ability, aptitude, personality, interests, etc., have scarcely been used. It may be noted that only one study has developed standardized entrance test. However, entrance tests for medical, engineering, IIT, and IIM are being currently used for admission purposes. Further, it has been observed that the examination marks have been very widely used for admission to different courses, for formative purposes for providing feedback to students, teachers, and system, and for summative purposes for certification. The educational institutions and employment agencies have used the same set of examination marks for the three above-mentioned purposes. This may be practical but it is definitely non-academic. It should be remembered that achievement tests or examination marks should be used for purposes of assessing the outcome of the instructional programmes. They cannot be used for predicting the success in schooling or later employments. The two objectives are different. Therefore, admission tests, entrance tests, and selection and promotion tests must be developed independently. Research in this direction is highly desira-

ble.

The problem of failures has been a very serious one as about one-half of the students fail in their final examinations. In other words, it means that one-half of the country's educational budget is wasted along with the added frustration in the youth and society. From a review of the researches in the area of failures it has been observed that no study has been conducted regarding the variations in cut-off scores and rate of failures in different school subjects and students. Of course, different cut-off scores for admissions have been provided for SC/ST students and also for students of reserved categories and also backward areas. Researchers should find out the relevance and procedure for developing different cut-off scores for different subjects, residence, sex, caste and region. The mastery of the subject at the primary stage and its impact on reducing failures at higher stages in different subject areas need attention of the researchers. Studies may be undertaken to test whether while marking answer books, the examiner is influenced by the fact that achievement is normally distributed.

The above issues, viz., achievement testing, diagnostic testing, examinations, prediction, and failures are always found to be centred around the traditional dimensions of achievement. The researchers in the area of achievement test construction should give due emphasis to the measurement of higher mental processes. Achievement test construction for the measurement of outcomes related to affective and psychomotor domains should be based on sound taxonomies of educational objectives in the affective and psychomotor domains. The trichotomy of domains is only a matter of analytical convenience. The objectives of different domains do interact with each other and in any given situation we are likely to come across an admixture of them. It has been noticed that affective objectives are put in the curriculum all right to start with but in course of time they tend to be eroded. Cognitive behaviours are more tangible and more easily observable than affective behaviours and, therefore, they are more amenable to teaching, evaluation, and research. But there is inadequacy of the tools of measurement in the affective domain, which demands the designing of tests for the evaluation of these objectives. The evaluation of the affective domain needs to be done as systematically as that of the cognitive and psychomotor domains. Achievement tests for affective and posychomotor domain have to be developed. The development of such tools has to be in consonance with the changed school activities and classroom teaching and learning situations.

The research and innovations in achievement testing and examinations need to be strengthened by introducing mechanical and electronic aids for designing, printing, sorting, scoring, and reporting, and also by strengthening institutions which are directly or indirectly connected with examinations and achievement testing. The use of mechanical aids like scanners (scoring machines) and computers would immediately demand the change of format, length and contents of the evaluation tools. Some boards of secondary education and other institutions of higher education have already attained some experience in this direction. Keeping in view such experiences, the researchers and administrators should design action research plans for the use of mechanical aids for improving the present evaluation system. In addition to the introduction of electronic gadgets and computers, a significant amount of funds and expert personnel are required to improve the examinations and achievement testing. The premier educational institutions like the UGC, the NCERT, the AIU and other institutions should further strengthen and coordinate their research and innovative roles in the area of educational evaluation and examinations.

ABSTRACTS : 1030 – 1082

1030. BHAT, R.N., GUNASEKARAN, K. and SHANMUGHAM, M., *A Comparative Study of the Grade Table and Direct Grading Methods*, Examination Reform Unit, Madras U., 1978

The major objectives of the investigation were: (i) to study the consistency of grade distribution within the examination and between the examiners using the two methods of evaluation, (ii) to study the discrepancy in the distribution of grades under both the methods of evaluation, (iii) to study the consistency within the examiner in evaluating individual questions/sections by using the two methods of evaluation, (iv) to study the extent of using the entire scale of evaluation, and (v) to study the opinions of the examiners about the two methods of evaluation.

The sample consisted of fifty answer scripts from among 397 scripts relating to the paper on Corporate Administration (including Company Law I) of the M.Com. examination of December 1976. Four examiners were selected from the list of examiners approved by the university for valuing papers in the subject. Each examiner had to value the scripts twice by each method with a time gap between each valuation. The same time gap between different valuations was attempted to be maintained for all the examiners. Each examiner, under each phase, was provided with a scheme of evaluation pertaining to the method (GTM/DGM), sheets for recording marks/grades and a copy of the question paper along with the answer scripts. A questionnaire was used to elicit opinions of the examiners regarding the practicability and suitability of the methods. Chi-square test was applied to test the association between the two sets of grades.

The major findings of the investigation were: (i) Three out of the four examiners were consistent within themselves in respect of grades under GTM. (ii) Each of the four examiners showed consistency within themselves in grading the scripts under DGM. (iii) The examiners were consistent between themselves when the scripts were ranked in order of merit under GTM as well as under DGM. (iv) There was no agreement between the examiners in respect of the distribution of grades by either method. (v) In general, there was more consistency in DGM evaluation than in GTM evaluation. (vi) Under GTM, the marks awarded by the four examiners ranged from zero to nine (maximum mark allotted to a question was twelve) but the grade points awarded under DGM ranged from zero to the maximum point. (vii) The examiners, in general, felt that objectivity in valuation was more in Direct Grading Method than in the Grade Table Method of evaluation. DGM consumed more time than GTM. They further opined that DGM might be adopted in the university after taking care to remove certain administrative problems.

***1031.** BHATNAGAR, T.N.S., *A Study of Some Cognitive and Motivational Factors in the Selected Population of the B.Ed. and Other Professional Courses (namely, Medical and Engineering) in relation to their Examination Success*, Ph.D. Edu., Mee. U., 1982

The specific objectives of the investigation were: (i) to study cognitive and motivational factors of student-teachers, medical students and engineering students, (ii) to compare student-teachers with medical students and engineering students on cognitive and motivational factors, (iii) to study the relationship between cognitive and motivational factors as well as achievement of the three groups of students, and (iv) to compare the extent of variation in respect of criterion variable through cognitive and motivational variables as predictors, among the three groups of students.

The sample for the study consisted of 281 students preparing for B.Ed. course, B.Tech. engineering courses and M.B.B.S. courses, drawn from the final year of these courses from the Central Institute of Education, Delhi, Indian Institute of Technology, New Delhi, and All India Institute of Medical Sciences, New Delhi. The three institutions were selected randomly through lottery system. The number of students selected was 167 (59.4 per cent) from education, 68 (42.0 per cent) from technology and 46 (16.4 per cent) from medicine. The tools for the study included eight tests of intellectual factors developed by French and others, Test of General Intelligence developed by Cattell, Abstract Reasoning Ability Test developed by Bennett and others, and Cattell's Motivation Analysis Test. Statistical techniques used for analysis of the data were mean, median, standard deviation, t-test, intercorrelations and step-wise regression analysis.

The major findings of the investigation were: (i) Student-teachers were similar to medical students in cognitive and motivational factors of general intelligence, abstract reasoning ability, home parental sentiment, super ego sentiment, etc., and were not at par with them

on certain other factors such as originality, semantic spontaneous flexibility, ideational fluency, career sentiment, self-concept sentiment, and so on. (ii) Student-teachers were similar to engineering students on abstract reasoning ability, home parental sentiment and were dissimilar on general intelligence, semantic spontaneous flexibility and career sentiment. (iii) Similarities were there between medical and engineering students in nearly all cognitive and motivational factors except one. (iv) In the case of student-teachers, maximum variance was explained by abstract reasoning ability followed by originality, pugnacity-sadism erg, home parental sentiment, semantic spontaneous flexibility, and so on. In the case of medical students, these were abstract reasoning ability followed by sweetheart spouse sentiment, assertive erg, and so on. Corresponding variables for engineering students were originality followed by semantic spontaneous flexibility, mating erg, and ideational fluency. (v) Abstract reasoning ability was the best predictor of achievement in the case of student-teachers and medical students. For engineering students, originality was the best predictor. (vi) One-fifth variance in the case of engineering and medical students and more than half in the case of student-teachers remained unexplained.

1032. BHATT, N.R. and INDIRESAN, J., *The Correlation of Performance of Students in High Schools with Their Achievement in Polytechnics,* TTTI, Madras, 1981

The objectives of the study were to find out (i) the usefulness and effectiveness of the present criteria in predicting the final achievement of students in polytechnics and (ii) the role of certain intervening factors like the socio-economic background of students in affecting their achievement in polytechnics. It was hypothesized that (i) there was no relationship between the performance in high school and the achievement in the diploma course and (ii) certain social factors like economic status, occupation of the father, social status, urban-rural background affected performance.

The study was conducted on 406 civil, mechanical and electrical engineering students who completed their diploma course in 1977 and were from five different polytechnics. Data were collected on their performance in English, mathematics, science and elective subjects in the high school examination, marks obtained in internal and external examinations in PTC semesters A, B, C and D (yielding ten sets of scores) and on income and occupation of father, community and rural-urban background. Correlation matrices were drawn for testing the relationship between the various scores.

The major findings of the study were: (i) The mean high school marks were significantly related to the scores in all the semesters and PTC in civil, mechanical and electrical engineering and both for internal and external assessments. (ii) The marks in history and geography were significantly related to all the sets of scores in mechanical engineering and all but two in electrical and civil engineering. (iii) The scores in science were significantly related to all but one set of scores in civil engineering (external of PTC), all but five sets in electrical and with all sets in mechanical. (iv) The mathematics scores were related to the civil engineering scores, five out of ten sets of scores in mechanical engineering and with seven sets of scores in electrical engineering. (v) There was no significant difference in the performance of the students belonging to different socio-economic categories. (vi) Though the high school marks were significantly related to the semester examination marks, they were not a good predictor of success in polytechnics and hence the high school scores as the only basis for admission needed to be reconsidered.

1033. BOSE, P.K., ROY, A. and MUKHERJEE, S.P., *The Existing System of Examinations and Measures for Improving upon It,* Dept. of Stat., Cal. U., 1967

The main aim of the study was to investigate the prevailing system of examinations and to suggest short-term remedial measures. An opinion survey was conducted during 1966-67. A questionnaire was circulated to teaching institutions and persons interested in education. The questionnaire included items on syllabus, the setting and moderation of question papers, conduct of examinations, admission to colleges, college tests, examination for external students, teaching in colleges, tutorial work, number of examinations and reports from colleges to the university. Opinions on all or some of the items were secured from 365 respondents comprising principals of affiliated colleges (N=48), teachers of affiliated colleges (N=266), teachers of Calcutta University (N=31) and other persons interested in education (N=20).

The findings of the study were: (i) On certain issues relating to the processes of learning and evaluation there was a consensus of opinion among the respondents, while on certain others their opinions differed. (ii) The college teachers and the principals branded the existing syllabi as too heavy. Most principals and some college

and university teachers wanted a change in the distribution of the total learning material between Part I and Part II. The existing syllabi in different subjects should be changed to ensure a smooth gradation from the syllabi in schools to the corresponding syllabi in university classes. (iii) A question paper should provide for double the number of questions to be answered and the questions should be distributed uniformly throughout the syllabus. (iv) The course material should be divided into groups and the question paper should be so set as to require the examinees to answer some questions from each group. Question papers should be moderated by the paper setters along with some other persons. (v) The university should undertake all the examinations in respect of collegiate education. Colleges should hold two or three periodical examinations in a year. Collegiate and external students should sit for the same examination. External candidates should be required to pass a test in some college in order to be eligible for appearing at the university examination. (vi) Opinions differed in respect of the method of teaching in colleges. (vii) The necessity for tutorial/seminar work was not accepted in the same spirit by all. (viii) Pass students should be required to pass Part I and Part II and Honours examinees be required to pass to their subsidiary subjects at these two examinations separately. (ix) Colleges should send reports of progress of study to the university once or twice a year. (x) Action should be taken against a college if the percentage of successful students sent by the college was poor for two or three consecutive years. (xi) An examination research unit should be permanently attached to the office of the Controller of Examinations.

1034. CHATTERJI, S. and MUKHERJEE, M., *Construction and Development of a Test of English Knowledge and Comprehension at the Higher Secondary Level,* Psychometric Research and Service Unit, ISI, Calcutta, 1970

The major objective was to develop and standardize a test for measuring knowledge and comprehension of English at the higher secondary level. The experimental version of the test was constructed with seventy-seven items forming two separately timed parts. Part I was for measuring English knowledge and it covered English usage, word meaning, grammar and spelling. Part II was designed for measuring comprehension of English, which was further subdivided into two parts. This version of the test was administered to 250 Class XI students reading in four different higher secondary schools in Calcutta and item analysis was conducted. The revised version of the test included sixty-seven items of which twenty-nine were in Part I and thirty-eight in Part II. The revised version was administered to 291 students studying in Class XI in five higher secondary schools as well as on 190 first year college students in two colleges at Calcutta. Reliability, validity and norms were established. Regression analysis was also carried out.

The major findings were: (i) Reliability of the two parts of the test was separately estimated by using K-R formula 21 for different groups and the values varied between 0.53 and 0.81. (ii) The validity coefficients, obtained by adopting different methods, for the two parts of the test varied widely from one instance to another. (iii) Correlation between the two parts was not very high, which indicated that the two parts of the test were not just duplicating each other. (iv) The test had quite a high predictive value.

1035. CHATTERJI, S. and MUKHERJEE, M., *Factorial Composition of School Examination Marks,* Psychometric Research and Service Unit, ISI, Calcutta, 1973

The main aim of the study was to extract the common factors present in school examination marks on different subjects. Twelve Bengali-medium schools at Calcutta were selected by using the random sampling method, after grouping them area-wise and sex-wise. Seven schools were boys' schools while the rest were girls' schools. From each school, the annual examination marks of Class VIII were collected covering English, Bengali, mathematics, history, geography, general science and Sanskrit. The analysis was done separately for each school. The intercorrelations among the subjects along with the means and standard deviations were calculated by using EDPM 1401 system. Each of the intercorrelation matrices was then factorized by using Thurstone's centroid method. Significance of the centroid factors was tested by Humphrey's rule and also by Burt's formula.

The findings of the study were: (i) Most of the school marks on all the seven subjects measured one general factor which was identified as the scholastic achievement factor. This suggested the presence of an halo effect in these cases. (ii) Only for a few schools two factors were extracted and in one case the second factor was identified as language factor.

1036. CHAUHAN, C.P.S., *An Evaluation of Achieve-*

ment in Algebra of Class IX Students in Delhi Schools with reference to Categories of Guilford's Structure of Intellect Model, Ph.D. Edu., JMI, 1982

The objectives of the study were: (i) to construct and standardize tests on eight selected structures of intellect as supposed to be involved in algebra, (ii) to construct and standardize eight tests in algebra on different steps necessary for algebraic problem-solving, (iii) to factor analyse the sixteen tests separately and also all the tests together, (iv) to conduct regression analysis of tests of structure of intellect abilities on tests in algebra, and (v) to determine the relationship between the results obtained from factor analysis and regression analysis.

For trying out the tests in algebra, 185 students of Class X were selected from Delhi schools and Kuder-Richardson Formula 21 was used to calculate the reliability coefficients of the tests. The validity of the tests in algebra was found out against the school marks in the subject. The eight algebra tests used in the study were Substitution Test (T1), Inequation Test (T2), Expression Rewriting Test (T3), Coefficient Comparison Test (T4), Definition Test (T5), Equation Formation Test (T6), Equivalent Equations Test (T7), and Sufficient Data Test (T8). The eight structure of intellect (S+I) tests were: C1, Letter Series (CSS); C2, Correlate Completion (NSR); C3, Word Transformation (NST); C4, Word Relations (CSR); C5, Synonyms (CMU); C6, Operations Sequence (NSS); C7, Word Patterns (CSI), and C8, Necessary Facts (CMS). Sixteen schools of Delhi were selected randomly. All the students of Class X in these schools comprised the sample of 500 students thus selected — 241 boys and 259 girls.

The findings of the study were: (i) Each of the achievement tests in algebra (T1 to T8) was significantly correlated with each of the SI tests (C1 to C8) for boys and girls separately as well as for the combined sample. (ii) Factor analysis of the tests of algebra indicated the existence of only one common factor, General Algebraic Reasoning, in the case of boys and girls separately and also in the combined sample. (iii) Factor analysis of the SI tests indicated only one common factor, Education of Correlates, characterized by C4 and C2, in the case of boys and in the case of girls two factors, namely, Education of Correlates and Verbal Comprehension. Only one factor was identified for the combined sample, namely, Education of Correlates. (iv) Two factors, namely, General Algebraic Reasoning and Pattern Visualization, were identified when all the sixteen tests including Ts and Cs were factor analysed for boys. Apart from the two factors identified in the case of boys, a third factor, General Reasoning, was also identified in the case of girls. In the case of combined sample, only two factors, General Algebraic Reasoning and Pattern Visualization, were identified. (v) Three SI abilities namely, Letter Series, Operations Sequence and Word Relations, were the best predictors of achievement on Substitution Test (T) in the case of boys, while for girls the predictors were the abilities Letter Series, Synonyms and Correlate Completion. For the combined sample two more abilities, Word Transformation and Necessary Facts, were predictors. (vi) For Inequation Test (T2), three abilities, Letter Series, Operations Sequence and Word Transformations, were identified as predictors in the case of boys. In the case of girls four abilities, Correlate Completion, Operations Sequence, Synonyms and Necessary Facts, and for the combined sample only three abilities, Letters Series, Operations Sequence and Correlate Completion, proved to be the best predictors. (vii) The best predictors of Expression Rewriting Test (T3) were CSS, NSS and CSR for boys, CSS, CSR and CMU for girls, and CSS, NSS, CSR, CMU and CMS for the combined sample. (viii) The predictors for Coefficient Comparison Test (T4) were CSS, NST, NSR and NSS for boys, CSS, CMU, NSR and NSS for girls, and CSS, NSR, CMU, NSS and NST for the combined sample. (ix) For Definitions Test (T5), the abilities identified as predictors were CSS and NST for boys, CSS, NST and CMS for girls and CSS, NST and CMU for the combined sample. (x) For Education Formation Test (T6), the predictors were CSS, NSS and CSR for boys, CSS, CMU and CSR for girls and CSS, NSS, CSR, CMU and CMS for the combined sample. (xi) Four tests, namely, CSS, NSS, NSR and CMS, acted as best predictors of achievement on Equivalent Equation Test (T7) for boys. For girls, four abilities, CSS, NST, CSR and CMU, were identified as predictors of T7. In the case of the combined sample five predictors, namely, CSS, CMS, NST, CSR and NSS, were identified. (xii) For Sufficient Data Test (T8), three SI abilities, CSS, NSS and CSR, were the best predictors in the case of boys. In the case of girls the best predictor abilities were CMU, CSS and CMS. For the combined sample five abilities, CSS, CMS, CMU, NSS and CSR, were identified as predictors.

1037. DE, U., *Development of a Scholastic Achievement Test in Geography for Madhyamik Students of West Bengal,* Ph.D. Edu., Cal. U., 1979

The main aim of the study was to develop a scholastic achievement test in geography for the madhyamik students of West Bengal. The final form of the test included nine different subtests, viz., classification, analogy, information, reasoning, comprehension, computation skill, sentence completion, analysis of relationship and map identification. The items were selected on the basis of difficulty values, discriminating powers and item-total correlations of the individual items. The tryout form and the final form of the test included 172 and 104 items, respectively. The standardization sample (N=2000) was selected by adopting the stratified random sampling technique, representing various geographical areas of West Bengal, which included boys and girls and covered different categories of schools up to the Class X stage. Reliability, validity and norms for the test were established by adopting different methods.

The findings of the study were: (i) The nature of distribution of the test scores, which were analysed statistically in three ways, was nearly normal. (ii) The reliability coefficient estimated by split-half method was 0.91. The reliability coefficient by K-R Formula 20 was 0.93. The reliability coefficients for each subtest obtained separately by K-R Formula 20 ranged between 0.45 and 0.91. (iii) The validity of the test was estimated by three ways and the test was found to be a valid one. Factor analysis results indicated three principal factors, viz., geographical reasoning, geographical aptitude and geographical application. (iv) Standard score, T-score, percentile and letter grade norms were developed for the total, boys and girls samples, separately. (v) Boys were superior to girls in general scholastic ability in geography.

1038. DEO, P., *Effects of Revaluation on the Results of Candidates Appearing at the University Examinations,* Dept. of Edu., Bom. U., 1980 (UGC-financed)

The main objectives of the investigation were: (i) to study the extent to which the marks awarded to candidates in university examinations changed on account of revaluation, (ii) to study the extent to which the results of candidates changed on account of revaluation, (iii) to examine the nature of the changes in the results due to revaluation, (iv) to find out the difference between the original and the revalued marks of candidates, (v) to examine the nature and size of differences in the scores between the original and the revalued marks, (vi) to find out the size of correlations between the original marks, the revalued marks and differences scores, (vii) to study whether the marginal differences acceptable to some universities were effective in changing the final marks and results, (viii) to study whether the different rules of revaluation, adopted by various universities, affected the results of candidates, and (ix) to make suggestions for improving the process of revaluation.

The design was an *ex post facto* design. Multi-stage cluster sampling was done; all the universities, excluding agricultural universities and technological institutes, were sampled. A proforma was prepared for the purpose. Data were collected for different university examinations for original marks, revalued marks, original and final results, university rules for revaluation regarding fees and rules for minimum difference for accepting a change in the result. Data were collected through personal visits to the universities. Statistical techniques such as mean, standard deviation, t-test, percentage difference and product moment correlation were used. The analysis of data was done by forming the following categories of candidates: (i) failed and already passed, (ii) change and no change in marks, (iii) change and no change in results, and (iv) marks increased and decreased cases. Similarly, the nature of the reassessing examiner *vis-a-vis* the original examiner was classified in the following categories: (i) consistently liberal, (ii) consistently strict, (iii) maintaining the same standard, and (iv) haphazard.

The main findings of the study were: (i) More students from the arts faculty applied for revaluation than from other faculties and more candidates who had failed applied for revaluation than those who had passed. (ii) Changes in results in the case of failed candidates were larger than no changes in marks. (iii) The proportion of the cases of increased marks was larger than the cases where there was decrease in marks. (iv) The proportion of marks-increased cases in the failed group was larger than that of marks-increased cases in the already-passed group. (v) The proportion of no changes in results was larger than changes in results. (vi) There was significant difference between the means of the original and the revalued marks. (vii) The correlation between the original and the revalued marks was high. (viii) The correlation of difference scores with the original and the revalued marks was low and sometimes negative. (ix) Most changes of revaluation cases were due to the marginal difference which were applied by the universities. (x) The high correlation between the original and the revalued marks established the reliability of assessment of essay-type questions. (xi) Small differences between the original and the revalued marks also contributed to

changes in the results of candidates. (xii) Revaluation created many administrative, psychological and financial problems.

1039. DIRECTORATE OF HIGHER EDUCATION, *A Study of Incidence of High Percentage of Failures in Public Examination of H.S.C.*, Hyderabad, 1966

The investigation attempted to study the reasons for the high percentage of failures in the public examination taken at the secondary stage in Andhra Pradesh. The sample consisted of twenty high schools from Andhra Pradesh. These schools produced very poor results in the H.S.C. Examination held in March/April 1963. A questionnaire was constructed to collect data on certain aspects from the institutions. It cosisted of eight items, both closed and open type. Interviews were held with the heads of the selected high schools. Frequency distributions, mean and percentage were employed to analyse the data.

The investigation yielded the following findings: (i) In Andhra Pradesh, out of 72,078 candidates who took their school final examination in March 1963, 29,380 candidates were declared to have passed. (ii) The different systems of examination were continuing in the State; high schools in nine Telangana districts sent candidates to the former high school certificate examination of the erstwhile State of Hyderabad and the eleven districts in the former Andhra region sent candidates at the end of eleventh class to the S.S.L.C. public examination. The curricula and the question papers differed in both the regions. (iii) The majority of the schools were originally primary schools which were later on raised to the status of high schools without any significant change either in accommodation or in equipment that should accompany the development of a primary school into a high school. (iv) Accommodation and furniture were inadequate and unsuitable in various schools. There was no library worth the name in any of the high schools taken up for the investigation. There was no laboratory worth the name and where a small laboratory existed it was not put to use at all by the teachers handling the science subjects. (v) In most of the schools, untrained teachers were working. Many of them were not adequately qualified to teach their subjects. In most cases the schools were understaffed and vacancies were not filled for a long period. (vi) The pupils in almost all the schools were from very poor families. Their home environment was not conducive to higher education. (vii) Most of the headmasters did not exercise any systematic supervision over the assistants handling the subjects. Most of the schools did not receive even a single visit by the inspecting officers of the education department. (viii) The lack of frequent visits by the officers of the education department deprived the members of the staff of an opportunity to get expert guidance in their work.

1040. GADGIL, A.V., *Study of the Causes of Large Failures in English at the S.S.C. Examination (Std. X) of March 1977*, Indian Institute of Education, Pune, 1978

The major aim of the study was to find out the causes for large incidence of failures in English at the S.S.C. Examination held in March 1977. It was hypothesized that the large incidence of failures was due to inadequate grounding in the subject, inadequate mastery of candidates over certain areas like spelling, inadequate mastery of teachers over the subject, inadequate coverage of the syllabus, expectation of insignificant text content knowledge and recall, lengthy question paper, inadequate motivation for study, inadequate attention paid to comprehension, composition and translation, inadequate guidance in the practice of the language and unnecessary copying of the question paper in the answer books.

The sample comprised — 29 schools, 26,924 candidates in English, and 100 answer scripts selected at random. A questionnaire was used as the tool for data collection. Further data were obtained from the records of the S.S.C.E. Board and analysis of answer scripts. The data were converted into percentages for analysis, comparison and interpretation.

The major findings of the study were: (i) About 11 per cent students secured less than ten marks and about 60 per cent secured less than twenty-four marks. This indicated unsatisfactory state of affairs in the teaching of English. (ii) The students were weak in translation, comprehension, letter-writing and composition. Even those questions which required answers in one sentence only were not answered satisfactorily. Questions in the areas of transformation, indirect type grammar and linguistics were answered satisfactorily. (iii) The students were weak in the use of articles, writing correct word order, combining sentences, use of phrases, discrimination of the correct word from pairs of similar words, reported speech, and use of verb forms. (iv) The students were poor in the comprehension and understanding of the

language. (v) Failure in English on a large scale was due to inadequate grounding in the subject, inadequate mastery of the candidates in certain areas of language learning, inadequacy of mastery of the teachers over the subject, inadequate coverage of the syllabus, inadequate attention paid to composition, comprehension and translation, inadequate motivation for study and inadequate guidance provided to students in the practice of the language.

1041. GADGIL, A.V., *Study of the Causes of Large Failures in Mathematics at the S.S.C. Examination (Std. X) of March 1977,* Indian Institute of Education, Pune, 1979

The inquiry was undertaken with a view to studying the causes of large incidence of failures in mathematics at the S.S.C. examination for Standard X held in March 1977. It was hypothesized that failure in mathematics was due to inadequate grounding in the subject, inadequate mastery of candiates over certain areas as problem-solving, graph, logarithms, inadequate mastery of teachers over the subject, inadequate coverage of the syllabus, inadequate time for solving the question paper, difficult questions in examination, inadequate motivation for the study of the subject and inadequate guidance in solving examples.

The sources of data were the mark-sheets of S.S.C. examination, and a sample of randomly selected answer scripts. The tool used was a questionnaire meant for schools. The data were converted into percentages. The sample comprised 27,990 pupils for algebra and 19,869 for geometry out of a total of 2,56,940.

The major findings of the inquiry were: (i) Results in algebra and geometry were comparable and it could not be concluded that more failures were due to more failures in algebra or in geometry. (ii) In algebra, about 37 per cent candidates secured less than ten marks out of seventy-five and about 67 per cent secured less than twenty-one marks. In geometry, about 27 per cent secured less than ten marks and about 71 per cent less than twenty-eight marks. The distribution was positively skewed in both the cases, indicating an unsatisfactory learning of the subjects. (iii) The sampled schools were small in size. About 46 per cent of the schools had a strength of fewer than 300 students and about 75 per cent had fewer than 500 students. Small schools were not economically viable and they had poor equipment, which led to poor performance by pupils. (iv) The percentage of the successful candidates in mathematics in 1977 was lower than that in 1976 and 1975. (v) Even though the percentage of trained graduate teachers in the twenty-nine schools included in the sample was satisfactory, only two teachers per school had taken some orientation courses in mathematics which could be considered satisfactory. (vi) Reasons for failure in mathematics were inadequate coverage of the syllabus, inadequate attention paid to some difficult topics, inadequate motivation for study and inadequate guidance provided to pupils for study.

1042. GADGIL, A.V., *Study of the Causes of Large Failures in Social Science at the S.S.C. Examination (Std. X) of March 1979,* Indian Institute of Education, Pune, 1979

The major objective of the study was to investigate the causes of failures in social science in the public examination at the end of Standard X held by the S.S.C.E. Board, Pune, in March, 1968. The major hypothesis examined was that failures in social science were due to (i) inadequate grounding in the subject; (ii) inadequate mastery of students over certain areas of the subject; (iii) inadequate mastery of teachers over the subject; (iv) inadequate time for the coverage of the subject; (v) questions of high difficulty level; (vi) unsatisfactory translation of the original question paper from English into other languages; (vii) inadequate motivation of students for study; and (viii) inadequate guidance in writing answers to questions.

Sources of data included a random sample of answer scripts and question papers. A questionnaire for the schools and a questionnaire for teachers formed the tools. The sample comprised 28,449 students from 55 examination centres out of 2,27,893 students for history and civics and 27,766 out of 2,27,893 students in geography.

The main conclusions of the study were: (i) Failures in social sciences were due to the reasons mentioned in the hypothesis. (ii) Results in social sciences did not show any improvement from 1975 onwards except in the year 1977. (iii) The distribution of marks gave a positively skewed curve indicating that the teaching-learning process in the subject was not satisfactory. (iv) The performance of the students was comparatively better in geography than in history and civics. (v) Civics was neglected by a large number of students. (vi) Question papers focussed on knowledge, information and memory testing. (vii) More than 36 per cent of the schools in the sample were small (with fewer than 300 students). These

schools were not viable units and were poor in equipment and hence poor in quality. (viii) In some cases, there was a sudden rise in the number of students in Class VIII indicating that the students who were not well prepared, were pushed into Class VIII from primary schools. (ix) The number of schools which did not provide adequate homework and did not check the same was not negligible. This factor also contributed to increased failures. (x) Though the percentage of trained teachers was satisfactory, the percentage of teachers qualified to teach social science was likely to be negligible. This also contributed to large-scale failures.

1043. GCPI, *An Experiment in Reforming the Examination System,* Allahabad, 1971

The objective of the study was: (i) to design such question papers as could check the students' tendency to cram, (ii) to develop comprehension and creative thinking abilities in students, and (iii) to prevent students from copying from books or notes in the examination hall.

The study was carried out on a sample of 102 students of Grade VIII, who were divided into six groups of fifteen to twenty students each for the purpose of monthly examinations held in September and January. Each of these groups worked under the supervision of separate student-teachers and each student-teacher was responsible for conducting and evaluating the examination for his group independently. Improved question papers were prepared for subjects like English and Hindi languages, mathematics, and general science. Students' achievement in different subjects was taken from the results of the six-monthly examinations conducted by the intermediate college. Each of the improved question papers for the monthly examination carrying twenty marks, was of thirty-five minutes' duration. The students were permitted to bring with them books, notes, etc., for copying in the examination hall. The improved question papers were based on the implementation aspect of the knowledge of various learned concepts. The questions based upon the direct or indirect reproduction of memorized facts, had no place in these question papers. The comparison of students' achievement in different subjects in monthly examinations of September and January and in the six-monthly examinations was made by computing the frequencies and percentages.

The analysis of the data revealed: (i) In comparison to the six-monthly examinations, the percentage of the students passing in the monthly examinations was very low, irrespective of the subjects. Few students obtained first class in the examinations. (ii) Even though the students had complete freedom to copy from books or notes during the monthly examinations, they could not score well in comparison to the six-monthly examinations where copying was not permitted. This shows that even copying in the examination hall will not help students, when question papers are set in an improved manner. (iii) The low pass percentage of the students in the monthly examinations indicated that only the really good students could pass with the improved question papers. (iv) The examination system based upon improved techniques of paper setting seemed to fulfil those aims and objectives of education where much emphasis was laid on original and logical thinking. (v) It was more difficult to set improved question papers in history than in mathematics, science, Hindi and English. (vi) For the successful implementation of the system of reformed question papers, it was recommended that teachers should change their teaching procedures accordingly, that the curriculum should be reorganized and that the evaluation system should be such that it evaluates all aspects of the child's personality — emotional, moral, skills, aptitudes, etc. — making daily, weekly and monthly evaluations necessary.

1044. GCPI, *A Commentary on the Evaluation Questions in the Lesson Plans of Mathematics and Science,* Allahabad, 1976

The objectives of the investigation were: (i) to find out the various teaching objectives that are likely to be fulfilled by the evaluation questions used by student-teachers, (ii) to find out the nature of the evaluation questions, (iii) to find out the average time required for answering these questions, and (iv) to suggest measures for the construction of proper evaluation questions.

One hundred and fifteen lesson plans prepared by the student-teachers for teaching the students from Grades VI to VIII were analysed to fulfil the objectives of the study. Of these, thirty-nine lesson plans were for mathematics, and thirty-eight each for science and general science. The analysis of these lesson plans was done keeping in view the objectives fulfilled by them.

The findings of the study were: (i) In the lesson plans pertaining to mathematics, the questions related to the knowledge aspect were given prime importance while the questions related to the application aspect were almost neglected. (ii) The knowledge-type questions were given maximum importance in the lesson plans for sci-

ence whereas the application-type questions were totally neglected. The questions testing skills were incorporated only in the biology lesson plans. (iii) The objective-type questions found much less place in the lesson plans for mathematics though such questions found an important place in the lesson plans for science. (iv) Two-thirds of the questions in mathematics as well as in science were of short-answer type. (v) The average expected time for completing the evaluation questions was about ten minutes.

Some of the suggestions to improve question papers were: (i) There should be a clear analysis of the objectives or the behavioural outcomes expected to be fulfilled by a particular lesson. (ii) The objective pertaining to the knowledge aspect should be divided into four parts — knowledge, understanding, skill and application. (iii) Under the objective pertaining to the evaluation aspect, higher mental abilities and skills like analysis, synthesis and evaluation and the like should also be given proper consideration. (iv) Questions should be set so that they challenge the abilities of students with higher mental processes. (v) Evaluation questions should not be restricted to the short-or long-answer type only.

1045. GCPI, *A Study of the Factors Responsible for Good Examination Results,* Allahabad, 1981

The objectives of the study were: (i) to find out the various factors responsible for increasing the percentage of examination results, (ii) to find out the various factors responsible for decreasing the percentage of examination results, (iii) to study the contribution of co-curricular activities in affecting the examination results, and (iv) to suggest possible measures to improve upon the examination results.

Ten schools from the city of Allahabad and its adjoining areas were selected on the basis of their examination results during the three consecutive years, 1977, 1978 and 1979. Of these, five schools were identified as having good percentage of examination results and the remaining five had poor percentage of examination results. Eight of these schools were for boys and two for girls; of these, seven schools were situated in the urban areas while three in the rural areas. Information was collected from ten principals and fifty teachers of these schools. The various instruments used for data collection were a school study proforma, a questionnaire for the principals and a questionnaire for the teachers. Data were analysed by computing the frequencies and percentages.

The findings of the study were: (i) A good school building, a good laboratory, good furniture, proper library and reading-room facility, playground, games and sports, appropriate situation and good environment of the school helped in improving the examination results. (ii) The teaching experience of the principal, capable and experienced staff, good methods of teaching, regular correction of homework, regular evaluation, proper attention to individual differences of students, proper educational guidance and encouragement to students, good academic achievement of the students at the time of admission to the school, good socio-economic status of the students, healthy relationship between the principal and the staff, proper cooperation between the teachers and the parents, good management and good discipline were the other factors which were significantly effective in improving the examination results. (iii) Discussion with teachers with respect to the various curricular activities to be covered during the session was also a significant factor in improving the results. (iv) There were no significant differences between those schools which had good examination results and those which had poor examination results with respect to the number of working days, the work-load of the teachers, the teacher-pupil ratio and the rules and regulations for admission and promotion of students. (v) Both the types of schools had an attitude of indifference towards the professional or academic growth of their teachers. (vi) The various factors responsible for poor examination results were the lack of dedicated teachers, indiscipline and the lack of interest in studies among students, passive attitude of parents towards the education of their wards, lack of proper correction of the homework, unnecessary interference of the members of students' union in the activities of the school, lack of material resources in the school, the copying and guessing tendencies of students, reading cheap and short-cut books, students' lack of interest in co-curricular activities, and the teachers' involvement in private tuitions.

The various measures recommended for the upliftment of examination results were increase in working days, provisions for diagnostic teaching, regular correction of homework, limited admission, provision for co-curricular activities, completion and revision of the course on time, restrictions on the publication of cheap books and guess papers, seeking cooperation of students and parents in the maintenance of discipline and encouragement to teachers to bring forth good examination results.

1046. GCPI, *An Evaluation Study of the Home Exami-*

nation System in Vogue, Allahabad, 1981

The objectives of the investigation were: (i) to compare the existing system of home examinations and the earlier one, (ii) to study the effectiveness of the present home examination system, (iii) to examine the appropriateness of the supplementary examinations, and (iv) to find out the various problems and their probable reasons that came in the way of the execution of home examinations and supplementary examinations.

The investigation was carried out in thirty secondary schools of which four schools were in the hilly areas, eleven in the western region, six in the central region, six in the eastern region and three in the Bundelkhand region of Uttar Pradesh. Information was collected from thirty principals of these schools, one hundred parents of children studying in these schools and five teachers, five students and five parents of the students studying in each of these thirty schools. Data were collected with the help of various questionnaires pertaining to the relevance of the existing home examination system and the earlier home examination system prepared separately for students, teachers, principals and parents, an analysis of the existing rules for passing the home and public examinations and the question papers set by the Board of Secondary Education for the years 1978-1980. The percentages and frequencies were computed for analysing the data.

The findings of the investigation were: (i) The supplementary examination should be abolished. (ii) The home examination should include evaluation of homework, classwork, co-curricular activities and discipline, at least 20 per cent weightage being given to these aspects. (iii) In all the regions of the State the home examinations should be conducted at almost the same time. (iv) There should be similarity in the course to be covered for the different home examinations throughout the State. (v) Each paper should be of two hundred marks of which forty marks be set aside for the first term examination, sixty marks for the six-monthly examination and one hundred marks be allotted for the annual examination. (vi) Question papers should be on the lines of the Board of Secondary Education Examination. (vii) Every student must appear in at least two examinations. (viii) Promotion to the next grade should be on the basis of aggregate of marks obtained by the student in three consecutive examinations, and the pass percentage be 33. (ix) In case a student failed in three subjects in Grades VI to VIII and in two subjects in Grades IX and XI he should be entitled to grace marks on the condition that the total aggregate of marks was not less than 33 per cent and the total amount of grace marks did not exceed fifteen marks.

1047. GUNASEKARAN, K. and JAYANTHI, P., *A Study to Fix the Permissible Difference in Grades Awarded by Two Examiners under Direct Grading Method,* Examination Reform Unit, Madras U., 1979

The major objective of the study was to fix the permissible difference in grades awarded by two examiners above which the scripts should be sent to the third examiner for valuation.

The procedure was to study the design of the question paper, the marks allotted to each question and the scheme for arriving at the grade point average and the grades. Three approaches to solve the problem were considered — a 10 per cent criteria, the permutation-combination method and the contribution method.

The conclusions of the study were: (i) The variations in terms of grades for each question might be found in terms of points also. (ii) The permissible difference between two examiners might be fixed in terms of the final grade arrived at for a script as it was not feasible to fix the permissible difference either in terms of each question or each section. (iii) The permissible difference between grades assigned by two examiners was fixed as one grade.

1048. GUNASEKARAN, K. and JAYANTHI, P., *A Study of the Continuous Internal Assessment and the University Examination Marks of the Undergraduate Semester Courses (1976-77 Batch),* Examination Reform Unit, Madras U., 1980

The major objectives of the study were: (i) to find out the correlation between the marks of continuous internal assessment (CIA) and the university examination (UE), (ii) to compare the marks in CIA and UE, (iii) to examine the distribution of marks in community social service (CSS), and (iv) to compare the distribution of divisions awarded in the non-semester and semester systems.

The study dealt with one paper from history, economics, philosophy, chemistry, botany, zoology and commerce. All the colleges offering these subjects were included in the study. Descriptive statistics and product moment coefficient of correlation were used for the analysis of the data.

The major findings of the study were: (i) The

minimum, maximum and the average marks awarded by the colleges in the internal assessment differed considerably in all the subjects. (ii) In the university examinations also, the minimum, maximum and average marks varied considerably among the colleges. (iii) Discrimination between good and poor students was low in the internal assessment but the university examinations showed a good amount of dispersion with regard to the same set of students. (iv) Barring a few cases, the relationship between the marks of the internal assessment and the university examination was good. (v) The assessment in the community social service varied among colleges. In most cases the assessment was on the higher side. In many cases maximum marks were awarded uniformly to entire groups of students. (vi) The percentage of those passing and those getting higher classes had improved under the semester system as compared to the non-semester system. More students secured first and second divisions under the semester system.

1049. GUNASEKARAN, K. and JAYANTHI, P., *A Study of the Revaluation Cases of the B.G.L. and B.L. Degree Examinations*, Examination Reform Unit, Madras U., 1981

The major objectives of the study were: (i) to find out the effect of revaluation on the final classification of students in terms of first, second and third divisions, (ii) to identify the examiners or a combination of them whose marking of the scripts consistently produced favourable or unfavourable results, (iii) to identify the errors committed by students and also the errors committed by examiners in valuing the scripts, and (iv) to identify the defects, if any, in the design of the question paper.

The sample consisted of all the cases who had applied for revaluation of scripts of examinations of B.G.L. and B.L. courses held during 1978 and 1979. They were 340 in the non-semester system and 31 in the semester system.

The findings of the study were: (i) Out of the 340 cases in the non-semester system, 81 candidates benefited as a result of revaluation. The corresponding number for the semester system was 13 out of the 31 cases. About 25 per cent applicants benefited as a result of revaluation. (ii) The maximum percentage of increase in marks as a result of revaluation was 35 and the corresponding percentage for lowering the marks was 13. (iii) In eighteen cases, the revaluation marks were non-beneficial to the candidates. (iv) In three cases, the candidates who had been declared as passing the examination, were declared as failing in the examination as a result of revaluation. (v) Certain inconsistencies were noted in the application of the rules for revaluation. (vi) As far as the award of classes was concerned, about 16 per cent of the applicants benefited. (vii) In the case of certain examiners, as many as twenty-one cases for revaluation were received and revaluation benefited twelve cases. (viii) The most common error committed by students was in giving wrong numbers to the questions. The common errors of the examiners were: (a) not awarding marks for all subdivisions of questions, (b) non-uniform award of marks to similar answers, (c) not awarding marks to the answers of all the questions, (d) awarding marks to more questions than required to be answered, (e) valuing the questions for more marks than the maximum. (ix) The question paper showed the defect of alloting fractional marks for each question. The question paper did not show the break-up of marks for each sub-division. The ambiguous wording of the questions was another common defect of the question paper.

1050. JOSHI, A.S., *Standardization of the Proficiency Test in Second Language for Senior Primary School Teachers of Gujarat State*, Ph.D. Edu., SPU, 1980

The main objective of the study was to construct and standardize second language (Hindi) proficiency test for senior primary school teachers of Gujarat.

There were 150 test items in the test. They were divided into three major areas — vocabulary, written expression and reading comprehension. The sample for the study consisting of 3,818 primary school teachers, was drawn from all the districts of Gujarat. Reliability was found out with the help of Rulon formula, Flanagan formula and K-R 20 formula. The reliability coefficient of the test as found by different methods was between 0.82 and 0.98. The comparison of reliability of the present test with another language ability test in second language showed that the present test had a very high reliability.

The content validity, construct validity, concurrent validity and predictive validity of the test were established. For the factorial validity three areas of the test were treated separately. The only significant factor in the area of vocabulary was correct meaning of the word. In the area of written expression the use of appropriate prefix and the effective use of punctuation marks were the significant factors. In the area of reading comprehension there were two significant factors. The re-

lationship between the language proficiency and the other factors was studied and reported in the thesis.

1051. KETKAR, S.R., *Construction and Standardization of Unit Tests in Mathematics for Standard VIII (Marathi Medium) based on the Syllabus of Maharashtra State,* Ph.D. Edu., Poona U., 1982

The major objectives of the study were: (i) to prepare unit tests covering the whole syllabus of mathematics for Class VIII, and (ii) to standardize the unit tests by finding out the difficulty value and the discriminating index of each item, determining the reliability and validity of the tests and fixing urban and rural norms for boys and girls.

The initial sample consisted of about 4,959 pupils from rural and urban areas from randomly selected schools of all talukas of Pune district. Standard procedures for writing test items, item analysis and standardization were followed.

The major conclusions of the study were: (i) The values of the mean went on decreasing from the first test to the last test indicating that the tests gradually became more difficult. (ii) In all the unit tests, except the first, the performance of the urban boys was superior to that of the rural boys. The urban girls had better performance than the rural girls on all the tests. (iii) The children found the algebra tests easier than the geometry tests. (iv) The distributions of the scores on all the tests were positively skewed. (v) The reliability coefficients ranged from 0.77 to 0.88.

1052. KHANDEWALE, S.S., *Construction and Standardization of Achievement Tests in Physics for Class IX in Vidarbha Region,* Ph.D. Edu., Nag. U., 1981

The investigation aimed at constructing and standardizing achievement tests in physics for Class IX. The tests were based on the syllabus of physics for Class IX in Maharashtra. Two sets of achievement tests were constructed. One set covered first five units of the syllabus and was to be used at the end of the first term. The second set was based on the remaining five units of the syllabus and was to be used at the end of the second term.

The tests were standardized on the population of the students of Class IX in the secondary schools of the Vidarbha region of Maharashtra. The sample for the tryout of the tests consisted of 200 students drawn from five schools of the region. The sample for the final normative study included 1,200 students drawn from thirty-three secondary schools. The tests were based on the prescribed syllabus and covering four objectives, namely, knowledge, understanding, application and skill. A two-dimensional blueprint was prepared giving weightage to different objectives and various subunits of the syllabus. Item analysis was carried out to find out the item statistics.

The tests were validated and their reliability was also established. The reliability coefficients using the split-half method and Spearman-Brown formula were found to be 0.85 and 0.84 for the two tests. Reliability coefficients as found out by using K-R formula 21 were 0.93 and 0.94 for the two tests. Indices of reliability were also calculated. The test scores were correlated with the examination marks to establish concurrent validity. The coefficients of criterion related validity (concurrent validity) for the two tests were 0.83 and 0.85. Percentile, stanine, Z-score and T-score norms were worked out.

1053. KOUL, L., *Scaling of Some Significant Innovations in Examination System,* School of Edu., HPU, 1979 (HPU-financed)

The major objective of the study was to investigate the views of university teachers and students regarding the relative scale positions of significant innovations and to compare the two sets of ratings.

Two samples of 250 postgraduate students and 80 teachers were selected randomly from the various faculties of the Postgraduate Centre and Directorate of Correspondence Courses of Himachal Pradesh University. The tool consisted of a checklist developed by the author from a list of thirteen innovations. Each one of the thirteen innovations was combined in pairs with every other innovation and thereby seventy-eight pairs were formed. The pairs were arranged and enlisted in such a manner that every innovation in the twelve pairs in which it appeared, occupied the first and second positions equal number of times and the pairs in which it appeared were kept as far apart as possible in the checklist. Thurstone's paired comparison method was applied to determine the scale positions of the innovations for the samples of teachers and students separately.

The major findings of the study were: (i) There was a good deal of similarity in the scale positions of the innovations according to the views of the teachers and the students. (ii) Certain innovations like introducing the grade system of ranking, supplementing essay-type

examination with objective-type examination, semester system and supplementing external examinations with periodic internal assessment, were in the top regions of the scales of the teachers and the students. Though individuals among the university teachers and students might differ in their views about the relative importance of innovations in the examination system, the collective opinion of both the groups was in favour of the grade system of ranking, internal assessment supplementing the external examinations, semester system and objective-type tests. (iii) The innovation, defining the scope of questions in simple and clear language, was also in the upper regions of both the scales. (iv) The teachers and the students differed regarding the scale position of the innovation, supplementing written examinations with oral tests. The teachers gave high importance to this innovation whereas the students gave low importance to oral tests. The university teachers were in agreement with the recommendation of the Education Commission (1964-66) but not the students. (v) Both the teachers and the students gave equal average importance to innovations such as orientation of paper-setters and evaluators and making evaluation instructions precise and clear. (vi) The innovations like open book examinations, use of computer and mechanical devices in preparing award rolls, scope for re-evaluation of answer books and spot evaluation of answer books, did not get importance from the teachers and the students.

1054. KUMAR, S.M., *A Study of Achievement of Kendriya Vidyalayas,* Ph.D. Edu., Pat. U., 1982

The aim of this evaluative study was to investigate the successes and failures of the Kendriya Vidyalayas, with respect to both academic and non-academic aspects. The sample comprised eighty-one respondents (thirty-four parents, twenty-seven teachers, ten administrators and ten educationists) randomly selected from all the eleven regions of the Kendriya Vidyalayas. A specially designed questionnaire was used covering areas such as the teaching at the primary level, language teaching, home assignment, time table, library facilities, examination results, admission rules, recreational activities, physical education activities, discipline, relationship between teachers and students, relationship between principals and students, relationship between teachers and parents, relationship between teachers and principal, parents meet, hostel facilities, sanitation, personal and social health, school environment and personality development of children. The investigator also examined the viability of the Kendriya Vidyalayas as model schools in respect of their organizational and administrative patterns, curriculum, school building and library and laboratory facilities.

The main findings of the study were: (i) The Kendriya Vidyalayas were functioning as model schools and as such the schools run and managed by other educational agencies in the country should follow the programmes of the Kendriya Vidyalayas. (ii) They were successfully inculcating among the students the attributes needed for national integration. (iii) They aimed at catering to the needs of civilian population and as such there was need for increasing the number of the Kendriya Vidyalayas. (iv) The curriculum and the textbooks of the Kendriya Vidyalayas were better suited for national integration. (v) The national pattern of education had been implemented in the true sense. (vi) There was need for improvement in teaching and library facilities.

1055. MASCARENHAS, M.J., *A Critical Survey of Examination Reforms Undertaken by the Maharashtra State Board of Secondary Education with special reference to the Question Papers Set in Higher Level English and in Geography (Special Geography and Geography in Social Studies),* Ph.D. Edu., Poona U., 1977

The major aims of the investigation were: (i) to survey the examination reforms undertaken by the Maharashtra State Board of Secondary Education, and (ii) to analyse the question papers in higher level English and in geography with the purpose of seeing their strength and weakness.

The tools and techniques used were questionnaires, discussions, interviews and scrutiny of question papers and answer books. Data were gathered from the heads of training colleges in Maharashtra, teachers of higher level English and geography, the officers of the S.S.C.E. Boards of Maharashtra, Gujarat and Andhra Pradesh, and the officers of the State Institute of Education of Andhra Pradesh. The sample of question papers scrutinized constituted ten question papers for higher level English and geography and five for social studies. Also, 510 answer books in higher English and 454 in geography were re-examined.

The major findings of the investigation were: (i) Between 1963 and 1965, six workshops were organized in Maharashtra which helped to propagate the evaluation approach among paper setters, examiners and moderators. After 1965, these ideas were found to have been

crystallized in the papers set by the divisional boards. (ii) The average pass percentage at the S.S.L.C. examination was about 30 prior to 1966, the average moved to 42 between 1966 and 1975. After 1975, there were four examinations (new course) with a pass percentage between 40 and 50 with the exception of October 1976 examination where the percentage dropped to 15. (iii) The syllabus in higher level English in the old and the new courses did not show any appreciable change. In geography, there was some significant change with the introduction of facts and figures in economic geography. The course in geography was full of factual information taxing the students' memory. (iv) A scrutiny of the past ten years' question papers in English and geography revealed domination of knowledge aspects with only a few items testing application or interest. The difficulty level had not changed. Most of the questions in geography had remained of the same type. A larger number of objective-type items had been introduced. However, the impression was that the paper appeared to be set for the below-average student. (v) In English, the standard was not adequate. Few students showed ability to discriminate, make inference or draw conclusions. Almost all questions needed recall rather than application. (vi) Only cognitive domain was being tested. (vii) No comprehensive scheme of internal assessment had been worked out.

1056. MITRA, S.K., CHATTERJEE, S. and MUKHERJEE, M., *Higher Secondary Science Achievement as related to Scientific Interest and Aptitude,* Psychometric Research and Service Unit, ISI, Calcutta, 1978

The main aim of the study was to find out the effect of scientific interest at different levels of potential ability with respect to achievement in science. The subjects were 115 boys studying in Class XI in three different schools at Calcutta, selected at random from Bengali-medium higher secondary boys' schools. The Scientific Knowledge and Aptitude Test and Chatterji's Non-Language Preference Record were administered. The higher secondary examination marks obtained by these students in the next year were collected. Relations among aptitude, interest and achievement were computed.

The findings of the study were: (i) Within the aptitude levels I and II, there was a high and very significant positive relationship between the intensity of scientific interest and the probabilities of success in higher secondary course. At the aptitude level III, however, a bit different result was observed; but as the number of cases at this level was very small such discrepancy might easily occur. (ii) Regression analysis results proved that the interest measure could significantly increase the efficiency of prediction of success when applied along with the aptitude scores. Scientific interest measure could be employed as a differentiating factor in the identification of more capable science students.

1057. MUZAFFAR, F., *Standardization of a Battery of Achievement Tests in Civics for Higher Secondary Stage in Madhya Pradesh,* Ph.D. Edu., Vik. U., 1967

The study aimed at standardizing a battery of achievement tests in civics for the higher secondary stage. For tryout, each test was administered on 370 students of one class. For fixing time limit each test was administered on 200 students of one class. The final form of each test was administered on 1,020 students of one class. The battery consisted of three tests each for Classes IX, X and XI. These tests were constructed on the basis of the course content prescribed for the Higher Secondary School Certificate Examination by the Board of Secondary Education, Madhya Pradesh, for the year 1964. The reliability of the tests was estimated by two methods — the split-half method and Kuder-Richardson formula. The empirical validity of the tests was worked out.

The findings of the study were: (i) The items having the difficulty value varying from thirty-five to sixty-five and the discriminative value varying between twenty and eighty-five were selected for the final tests. (ii) The reliability of the tests was estimated by two methods—by the split-half method, it was 0.99 for the three tests and by Kuder-Richardson formula it ranged from 0.91 to 0.92 for the three tests. (iii) The empirical validity of the tests ranged from 0.49 to 0.54. (iv) Percentiles, T-scores and letter grades were assigned for interpreting the scores on the tests.

1058. NAIR, J.R., *A Statistical Analysis of the Examination Results of the South Indian Universities,* Ker. U., 1982 (UGC-financed)

The study aimed to analyse and compare the examination results of five South Indian universities—Kerala, Calicut, Madras, Mysore and Tirupati. The S.S.L.C. Examination results of Kerala and Tamil Nadu, the pre-

degree/pre-university examination results of Calicut, Kerala and Madras Universities and B.A. and B.Sc. examination results of all the five universities were compared.

The details of S.S.L.C. Examination were collected from the office of the Secretary of Government Examinations, of the pre-degree and pre-university examinations from the records of the concerned universities and of the B.A. and B.Sc. examinations from Form B (Statistics) of universities collected for the University Grants Commission. The analysis of variance technique was applied to measure the variance between years, and the technique of multiple comparison was applied for certain situations.

The main conclusions of the study were: (i) There was significant variation in the percentage of passes during the years under study, for all the examinations studied, but the extent of variation differed among the universities. An important reason identified for the variation was the standard and difficulty level of the questions. (ii) Sex differences in achievement were identified for all the five universities with women showing higher achievement. (iii) The interaction between the sex and the class or division showed a differential pattern indicating significant difference for certain examinations of particular universities only. (iv) The forward community students were significantly better in pass percentage than the backward community students in Kerala.

1059. NARAYANA, T.V., *The Study of Performance Levels of Candidates Appearing at the Andhra Pradesh Public Service Commission Examinations Conducted during 1962-78 in relation to Candidates' Caste, Age and Education Background*, Ph.D. Edu., Kar. U., 1979

The main objectives of the study were to find out the relationship between (i) the levels of caste groups and their performance levels at the Public Service Commission examinations, (ii) the levels of sub-caste groups and their performance levels at the examinations, (iii) the levels of sub-caste groups and their performance level, (iv) the changing levels of performance of various caste groups, (v) the merit lists and the selection lists prepared on the basis of reservations, (vi) the educational levels of the candidates and their performance levels at the examinations, (vii) the number of positions secured by postgraduates and the number of positions secured by graduates, (viii) the number of positions secured by candidates with arts background and the number of positions secured by candidates with science background, (ix) the number of positions secured by the students who had secured merit in the university examinations and the number of positions secured by the students who had not secured merit in the university examinations, and (x) the age levels of candidates and their performance levels at the Andhra Pradesh Public Service Commission Examinations.

The sample consisted of the total population that was called for the oral tests during the years 1962 and 1978. For the purpose of comparison, F-values for performance, caste, subcaste, education and age were computed and t-test was applied for the two performance levels, caste-wise.

The significant findings of the investigation were: (i) The performance levels of the candidates appearing at the Andhra Pradesh Publice Service Commission examination increased with the caste levels. (ii) The performance levels of the candidates increased with the subcaste levels of the candidates barring the performance levels of the backward class. (iii) The oral performance level of the candidates increased with the caste levels and that of the sub-castes did not totally correspond with the caste hierarchy. (iv) The performance levels of the various caste groups improved over the years 1962-63 to 1976-78; the performance levels of the other caste group improved at a faster rate than the performance levels of the backward caste group and the performance level of backward caste group improved at a faster rate than the performance levels of the scheduled caste and scheduled tribe groups. (v) Because of the implementation of the rules relating to caste reservations, some groups gained additional selections during 1975 and 1978. (vi) The performance levels of the other caste group showed a slow decline with age; the performance levels of the backward class group showed a declining trend with age, from the 18-20 age group to the 24-26 age group, and an increasing trend between the age groups 24-26 to 30 and above; the performance levels of the scheduled caste and scheduled tribe group increased with age. (vii) The performance levels of the candidates increased with the levels of education up to seventeen years of education. The candidates with eighteen years of education, however, recorded lower levels of performance than the candidates with seventeen years of education. (viii) The second degree holders outnumbered the first degree holders in the selection of Group I and II services while the first degree holders outnumbered the second degree holders in the selection of Group III services. (ix) Almost the same percentage of candidates with arts and science were selected. (x) In the selection of all the three ser-

vices, viz., Group I, II and III, the candidates with first division predominated, and the candidates with second division predominated those with third division.

1060. NATH, B., *University Examination — an Analytical Study of the Conduct of Pre-university Degree and Master Degree Examinations of Gauhati University*, Ph.D. Edu., Gau. U., 1980

The main aim of the study was to present a systematic analysis of the prevailing system of essay-type examinations with special reference to the process of conduct and anlysis of results. The major aspects of examinations, viz., the question paper, mode of evaluation, scaling of marks and their impact on examination results, improvement of the design, structure and analysis of question papers in terms of difficulty values and discriminating powers of the items, the problem of optional questions in a question paper, desirability and methods of introduction of internal assessment as a tool for continuous evaluation, predictive value of an examination and wastage and stagnation, were studied systematically.

The findings of the study were: (i) The quality of students admitted and the teacher-pupil ratio had a bearing on examination results. (ii) There were large variations in the percentage of attempts not only in different question items but also in a number of sub-questions of the selected question paper. (iii) In a majority of the subjects, the scores were higher in internal assessment. The assessments, both in internal and external examinations, were evenly distributed. Internal marks possessed some predictive values for external marks. There was a tendency towards overmarking in internal assessment. Internal assessment provided some incentives to students to do better in the external examination. The system of internal assessment should be introduced, after fulfilment of certain conditions, for making the examination more valid and for doing justice to students. (iv) The rate of wastage and stagnation varied from subject to subject. The overall rate of wastage was slightly higher among males than among females. But the rate of stagnation among females was significantly higher than among males. The highest and the lowest rates of wastage were in English and botany, respectively. There were year-to-year variations in the rates of wastage and stagnation. The major causes of wastage and stagnation were poor economic condition of parents, admission of poor quality students, absence of proper system of internal assessment and continuous evaluation, and general apathy of students towards their courses of studies. (v) The performance of students in M.A. previous and final examinations in all the subjects was closely related.

1061. RAMACHANDRAN, A., CHATTERJEE, S. and MUKHERJEE, M., *Prediction of Scholastic Achievement in Two Major Streams of Study through a Multiple Discriminant Function Based on Interest Scores*, Psychometric Research and Service Unit, ISI, Calcutta, 1971

The major objective of the study was to construct a discriminant function which would guide in predicting scholastic achievement of higher secondary students in the two streams of study, viz., humanities and science, on the basis of interest scores.

The study included two different samples, viz., experimental (N = 145) and cross validation (N = 86). The experimental sample was drawn from six higher secondary girls' schools (seventy were from the humanities and seventy-five from science). The second sample of students used for the purpose of cross validation consisted of eighty-six girl students studying in Class XI in two different schools. Of these, fifty students were in the humanities group while the rest were in the science group. Chatterjee's Non-Language Preference Record was used.

The findings of the study were: (i) The interest patterns of the successful students in the science and the humanities streams differed widely. (ii) The students could be classified with considerable precision into science and humanities streams on the basis of their interest patterns.

1062. RASOOL, G., SARUP, R. and SHARMA, N.R., *A Comparative Study of Internal and External Awards at the Postgraduate Level in Jammu University*, Jammu U., 1981

The study primarily aimed at making a statistical analysis of the marks awarded by the external examiners and the marks awarded against the sessional work, i.e., internal assessment. The specific aims of the study were: (i) to examine the characteristics of the distribution of the scores awarded by the external and the internal examiners, (ii) to find out the degree of relationship between the external and the internal marks, and (iii) to find out the effect of the internal marks on the boosting of the overall result of students. The study was confined to 218 postgraduate students, both boys and girls,

selected from thirteen postgraduate teaching departments of Jammu University.

The major findings of the study were: (i) The scores in most of the papers showed deviation from the normal distribution. They were either positively or negatively skewed with a slight kurtosis. (ii) The majority of the teachers preferred to maintain a low range in the marks of the internal assessment, that is, the gap between the lowest and the highest scores was narrow. The range of the external marks in the same papers was definitely more than that of the internal marks. (iii) Most of the coefficients of correlation appeared to be positive. This tendency indicated that there was some conformity in the scoring pattern of the internal and the external examiners. (iv) The internal assessment proved to be a booster of the final result of almost all the students. The marks awarded under the internal assessment helped the students in raising their aggregate percentage of marks. Though the internal assessment suffered from various drawbacks like the halo effect and error due to central tendency, etc., it appeared to be a blessing in disguise to students.

1063. RAWAT, M.S., *Construction and Standardization of Diagnostic Test in Chemistry*, Ph.D. Edu., Agra U., 1976

The study was an attempt to construct and standardize a diagnostic test in chemistry for the intermediate stage in the State of U.P. The main objectives of the study were: (i) to construct a diagnostic test in chemistry (physical and organic), (ii) to ascertain the reliability and validity of the test, (iii) to analyse the errors committed by pupils and find out their frequencies, (iv) to categorize the errors, (v) to assign causes for the errors, and (vi) to prepare a remedial programme.

The sample of the study consisted of 600 boys and girls from rural and urban population randomly selected. Only twelve topics were selected from the field of chemistry (organic and physical) from the U.P. Board's syllabus on the basis of their high frequency in the examination papers. The number of teachers and experts from whom assistance was sought was fifty. Eight hundred answer scripts of intermediate students were analysed for the construction of the test. The objectives of knowledge, understanding and application were considered. The forms of the items were completion, alternate response, short answer and multiple choice. The test was tried out on a sample of 100 candidates and its reliability and validity were ascertained. The reliability of the test was 0.72 and the content validity was ascertained by comparing the errors through different sources like opinion, analysis of answer scripts of the pupils and the test itself. The responses under each item were analysed and discussed with experts in chemistry and then categorized. Then causes were assigned to each category of errors. The categories of errors were: (i) lack of knowledge of facts, concepts, process, new information, and definitions, (ii) errors in translation and in representing formulae and equations, (iii) errors in balancing the equation, (iv) errors in identification, (v) errors in discrimination, (vi) errors in comprehension, (vii) errors in establishing relationship, (viii) errors in the selection of correct information, (ix) errors in the application of proper formula, principles and processes, and (x) errors in computation. Apart from this a few remedial programmes were prepared and applied. The results showed the effectiveness of these programmes.

The major findings of the study were: (i) The fundamentals of chemistry like concepts, processes, principles and new information and discoveries were not clear to the pupils. (ii) The test results revealed that the pupils were not able to comprehend the fundamentals and translate the language of chemistry into symbols, formulae, equations, etc. (iii) The pupils were wanting in the interpretation of information. (iv) The knowledge of different units, their differences and conversions were not clear to the students. (v) The pupils were wanting in correctly balancing chemical equations and conversions of energy and chemical reactions. (vi) The ability to apply the knowledge of processes, principles and formulae was to be developed in the pupils.

1064. REDDY, A.V.R., *Attitude of Teachers of Two Institutions towards Internal Assessment — a Comparative Study*, Dept. of Edu., SVU, 1979

The investigation was designed to compare the views of the teachers of a university college where internal assessment (IA) was going to be introduced from the next academic year with those of the teachers of another university college where the system was being tried out for a few years.

The sample comprised 120 teachers equally distributed among the arts and science faculties in two university colleges in one of which (C1) IA was in practice for some years and in the other (C2) IA was to be introduced soon. Both the colleges were in the same university and had postgraduate courses. The sample included professors, readers and lecturers having at least two

years of experience in the profession. The research tool was a Likert-type attitude scale developed by the investigator, consisting of forty items (nineteen positive and twenty-one negative). The scale had a split-half reliability of 0.96 (N=50). The t-test and chi-square were the statistical techniques used for data analysis.

The major findings of the investigation were: (i) All the subgroups of the teachers except those belonging to the arts faculty of college C1 were significantly favourable towards their system of internal assessment. (ii) The teachers of the college where internal assessment had not yet started did not differ in their attitude towards the system from those of the college where the system was in practice. (iii) The teachers of the arts faculty of college C1 did not differ significantly from their counterparts in college C2 as regards their attitude to the system of internal assessment. The same was true of the teachers of the science faculty. (iv) Many teachers preferred partial IA with an external examination at the end of the year. (v) Discussion with the students and the teachers of college C1 revealed that the students depended mostly on lecture notes. This was also seen when more than half of the teachers felt that the students did not learn the subject better in this system. (vi) The teachers of college C1 felt that high hopes about the system of IA were belied when the system was implemented in their college.

1065. SALI, V.Z., *Question-wise Analysis of Answer Books of Mathematics and English of the Secondary School Leaving Examination: March 1977,* S.I.E., Pune, 1978

The objectives of the study were: (i) to analyse students' achievement in each constituent question in mathematics, (ii) to find out the general errors students committed while answering questions in mathematics, and (iii) to find out the drawbacks in the answer books of students.

The mathematics answer books of 300 students appearing for the school leaving examination of the Maharashtra Board of Secondary Education, Nagpur Division, of the year 1977, were selected for study. The random sampling technique was applied for choosing the answer books.

The study revealed: (i) In algebra, out of forty-five subitems in nine questions only three subitems were answered by more than 80 per cent students, fourteen subitems by only 10 to 20 per cent students and the remaining items by 20 to 80 per cent students. (ii) In one subitem 55 per cent students secured pass marks, in five subitems 40 to 50 per cent secured pass marks, in twenty-five items a maximum of 20 per cent students got pass marks, and in the case of fourteen items the pass percentage varied from 20 to 40. (iii) The errors in algebra were errors in transposition, forming the equations, solving equations with co-efficients and fractions, adding fractions, finding the antilog, assigning signs while factorizing, finding out LCM and HCF and finding ratio and proportion. (iv) The common errors found in geometry were errors in multiplication, writing figures, use of theorem on perpendicular bisector of chord, applying Pythagoras theorem, writing properties of isosceles triangle, naming the angle, finding the height of triangle, formula and calculation of area of segment of a circle, writing formula and putting values, measuring sides, constructing inscribed circles, using geometrical instruments, writing theorems on alternate angles and opposite angles. There were many other limitations, namely, in construction and determining the length of the base of a parallelogram, distinction between a parallelogram and a rectangle, using the formulae for finding out volumes. They also lacked in the knowledge of right angles of the base of a solid, properties and theorem of cyclic quadrilateral, properties of similar triangles, congruency of triangles, and geometrical means.

1066. SALI, V.Z. and UMATHE, V.T., *A Study of Relationship between Marks Obtained by Students in Theory and Practical Examinations in Science at S.S.C. Examination, 1977,* Maharashtra State Board of Secondary and Higher Secondary Education, Pune, 1979

The study aimed at investigating the relationship between the marks obtained by students in theory and practical examinations in science at the S.S.C. Examination, 1977, and suggesting a reliable marking scheme in practical examination in science. The hypothesis of no significant relationship between the theory and practical examination marks was formulated.

The sample comprised 800 pupils selected on a random basis out of about 80,000 students who appeared in the examination in the Vidarbha region of Maharashtra. Out of these 800 pupils, 626 had appeared in theory and practical examinations and their marks constituted the data.

The major findings of the study were: (i) Out of the maximum possible score of 120, the mean score of a candidate in science theory was 38.51 (32.10 per cent) whereas the mean score of a candidate in science practi-

cals was 21.61 (72.00 per cent) out of the maximum possible score of thirty. (ii) About 33 per cent students scored more than 80 per cent in practicals whereas only 17.5 per cent students scored less than 60 per cent marks in practicals. About 83 per cent students secured more than 60 per cent marks and about 4 per cent students secured 100 per cent marks. (iii) The results indicated the non-discriminatory nature of examination between good and poor students. (iv) The product moment coefficient of correlation between the marks in theory and practicals was 0.34, which was significant at 0.01 level. (v) Out of 626 students, 123 who secured less than 20 per cent marks in theory had a mean score of 66.4 per cent in the practicals, whereas 204 who secured more than 80 per cent marks in practicals had a mean score of 39.1 per cent in theory. (vi) The majority of the students (75 per cent) scored less than 40 per cent marks in theory whereas a majority of them (82.5 per cent) secured more than 60 per cent marks in practicals. Again, only 6 per cent students secured more than 60 per cent marks in theory against only 17.5 per cent pupils securing less than 60 per cent marks.

1067. SHARMA, A.K., *A Critical Study of the Achievement in Mathematics by Pupils of Secondary Schools with particular reference to the State of Assam,* Ph.D. Edu., Gau. U., 1978

The main objective of the study was to find out and assess the achievement in learning school mathematics and the possible reasons for the low achievement of the same. The study was confined to the areas of arithmetic and algebra of school mathematics.

A battery of sequential achievement tests in mathematics was constructed for Classes V to X. The items for each test were prepared from each topic of the respective syllabus covering such aspects as knowledge, understanding, skill and application. The draft tests were tried out. Based on the difficulty value and discriminating index of each item, 254 items (41, 42, 43, 48, 46 and 34 items for Classes V, VI, VII, VIII, IX and X, respectively) were selected in the final achievement tests. The reliability and validity coefficients of the tests were computed. The sample included 1295 pupils (210, 186, 178, 312, 263 and 146 from Classes V, VI, VII, VIII, IX and X, respectively) from ten schools. The mathematics teachers of these schools were interviewed with the help of a questionnaire. An analysis of syllabus, textbooks, exercise books, school records and board's records was also done.

The major findings of the study were: (i) The reliability coefficients of the achievement test battery obtained by the method of rational equivalence ranged from 0.66 to 0.75. (ii) The validity coefficients varied from 0.43 to 0.90 when correlated with the scores obtained by the same set of pupils in the school annual examination. (iii) The performance of the boys was better than that of the girls. (iv) The pupils of Classes V, VI and VII showed better results in the sub-classes of knowledge and skill than of understanding and application. (v) The major defects were the lack of drilling and knowledge on fundamentals and the inability to transform verbal statements into mathematical statements. (vi) The pupils of Classes VIII, IX and X were better in knowledge and skill than in understanding and application. (vii) The inadequate learning of the basic operations of elementary arithmetic at the primary stage led to difficulty in learning algebra at the secondary stage. (viii) Since all the pupils acquired knowledge and skill better than understanding and application of different topics, there was undue emphasis on the mechanical learning of mathematics. (ix) The pupils in urban, rural and backward areas did not differ very significantly in their performance, however, the performance of the rural boys was better than that of boys from the urban as well as the backward areas. (x) The urban girls showed better performance in understanding and application while the rural girls did better in knowledge and skill. (xi) The pupils of Classes VIII, IX and X performed better on the topics taught earlier in a session than the topics taught during the latter part. (xii) The coefficients of correlation between the performance of the pupils on topics in arithmetic and algebra in Classes VIII, IX and X showed the existence of a significant relationship for Classes VIII and X only. (xiii) Some major factors responsible for low achievement in mathematics were the imparting of limited knowledge, blind use of rules, heavy syllabus, defective textbooks, lack of the natural urge among pupils to learn mathematics, insufficient drill work at the primary stage and absence of methodical approach of the classroom teaching.

1068. SCERT, *Standardization of a Science Attainment Test for Class VIII in Telugu,* Hyderabad, 1971

The aim of the study was to construct and standardize a science attainment test for students of Standard VIII of Telugu medium. A large number of items were constructed in Telugu based on the science syllabus of Standard VIII of the high schools of Andhra Pradesh. Items

were also constructed for the new units incorporated in the science syllabus of Standard VIII. These items were tested on a small sample of Standard VIII students. Some of the items were dropped and a few were refined after the preliminary testing. The final version of the test was administered on 1,288 boys and 403 girls of Standard VIII from Andhra Predesh. Boys and girls were selected randomly from the various high schools of Andhra Pradesh.

The standardized test consisted of 100 questions. It was divided into two parts. Part I consisted of questions regarding physics and chemistry and Part II of questions regarding zoology and botany. Each part consisted of fifty questions. Each question had four answers and the students were supposed to find the correct answer out of the four answers.

The mean scores of Part I for boys and girls were 21.0 and 19.0, respectively. The critical ratio for Part I was 6.33. The average mean scores of Part II for boys and girls were 22.0 and 21.0, respectively. The critical ratio for Part II was 2.83. The mean scores of the complete test for boys and girls were 22.0 and 20.0, respectively. The critical ratio for the complete test was 8.29. The time required for administering the complete test was one hour and seven minutes. Part I required thirty-five minutes and Part II thirty-two minutes. Scoring key was developed. The test could be used as a yardstick to assess the knowledge of Standard VIII boys and girls of Telugu medium regarding their science attainment. Percentile scores were established for boys and girls.

1069. SCERT, *A Study on the Causes for the Poor Results of Class VII Common Examination held in March 1976,* Hyderabad, 1977

The investigation attempted to study the results of Class VII common examination of March 1976 from different points of view including the pattern of question papers, the scheme of evaluation, the teaching methods, the academic performance of the pupils in the previous classes and the instructional facilities in schools.

The sample consisted of a cross-section of the population involved in the process of education. Teachers, headmasters, inspecting officers, parents and students were involved in the same to give representation for their roles. In selecting schools, care was taken to involve schools of both boys and girls under different managements (viz., the government, the zilla parishad, the municipality and private). Warangal, Srikakulam and Kurnool districts and the twin cities of Hyderabad and Secunderabad constituted the sample area in Andhra Pradesh for study. Data were collected partly by mailing questionnaires to the respondents and partly through interview schedules. Opinions of the public were also invited through a press note. In all thirteen schedules were prepared. Four of them were questionnaires, three of them were case studies while six of them were structured interview schedules. The data through Schedules I to VII were collected through the district education officers of the three districts (Srikakulam, Kurnool and Warangal) and the twin cities. The data through Schedules VIII to XIII were collected by the investigators themselves. Percentage and chi-square were employed to analyse the data.

The investigation yielded the following findings: (i) There was no staff room and no books for general and supplementary reading for the teachers in nearly 60 per cent of the schools. (ii) In 11 per cent of the schools even a copy of the syllabus was not available. (iii) Instructional facilities in all the subjects were not adequate in the case of at least one half of the schools, though the position in the case of science and mathematics appeared to be slightly better than in the case of social studies. (iv) More than 90 per cent of the respondents felt that they needed preparation before teaching. (v) More than 86 per cent of the teachers professed that they had the skills of identifying facts, principles, etc., from the content, arranging them in a hierarchical order of difficulty. (vi) Approximately 80 per cent of the teachers professed that they were following the syllabus scrupulously in picking up objectives, in planning activities, in devising teaching steps in mathematics, in stressing the desirable vocabulary and in developing language skills. (vii) Reasons given by the students for the large number of failures were inability to understand what they studied, inability to follow the lessons taught in the classroom, the lack of practice in answering questions and study habits. (viii) Nearly 86 per cent of the examiners felt that the principles of valuation supplied to them were useful. (ix) About 37 per cent of them felt that they could not do justice to students by following the principles of valuation in valuing the scripts. (x) Nearly 82 per cent of the examiners felt that the essay type questions allowed greater discretion for them while 21.2 per cent said so in the case of short-answer questions and 17.6 per cent said so in the case of objective-type questions.

1070. SETHIA, K.K., *Diagnostic Testing and Remedial Teaching in Physics in Class IX,* Govt. Higher Secondary School, Deorakalan, 1972 (NCERT-financed)

The objectives of the study were: (i) to diagnose the nature and causes of unsatisfactory adjustment to school situations, (ii) to diagnose the nature of maladjustment for the individual learner and for the whole group, (iii) to diagnose the causes behind such maladjustments in every teaching unit in physics, (iv) to diagnose the causes which interfere with the normal academic progress of the individual, (v) to choose and apply remedies, and (vi) to evaluate the work before and after the remedial teaching.

This was an experimental study, experiment having been conducted on ten students of Class IX of the Government Higher Secondary School, Deorakalan. The students were of very low socio-economic order. They were administered thirty different diagnostic tests, one after each teaching. The thirty tests comprised eleven on general physics (properties), five on heat, seven on light, four on magnetism and three on electricity. After each diagnostic test remedial teaching was carried out, which consisted of personal guidance, group discussion and teaching in group, discussion of test papers, demonstration (by the teacher and students), preparation and use of charts and models, seeking cooperation of the family members and providing general help. The actual remedial programme was designed on the assessment of the needs of the individuals assessed through the diagnostic tests and personal interviews. The impact of the remedial programme was estimated by examining the results in the next quarterly examination and comparing this with the results of the diagnostic tests. The data were analysed in terms of individual gains and group gains in each test, each cluster and the whole test.

The findings of the study were: (i) In all the clusters—general, heat, light, magnetism and electricity—every student gained positively in the scores. The gain in each test ranged between 8 and 33 per cent. (ii) As a group, the gain was 13 per cent in general, 10 per cent in heat, 11.5 per cent in light, 14.5 per cent in magnetism and 13.2 per cent in electricity, between the diagnostic test and the quarterly test.

1071. SHAH, J.H., *Construction and Standardization of Primary School Achievement Tests (PSAT) for Pupils of Grade VII in the State of Gujarat,* Dept. of Edu., Guj. U, 1982 (NCERT-financed)

The aim of the investigation was to develop a battery of tests which could help to estimate the developed ability of seventh graders of Gujarat.

The PSAT consisted of four tests, which were vocabulary, routine computation, sentence completion and mathematical reasoning. Each test had twenty-five multiple choice items, thus having 100 items in all. It took two class periods for its administration, and yielded V score, Q score and total score. Three tryouts were carried out on eighty, two hundred, and three hundred and eighty-nine pupils to refine and select the items from the pool of 265 items for the final run. The final form was administered to 2,089 pupils (1,300 boys and 789 girls) drawn from fifty-two schools of different regions of Gujarat, by the stratified cluster sampling method. Stanine norms for V, Q and total scores were developed for three different groups, namely, (i) urban (boys and girls), (ii) semi-urban and rural (boys only) and (iii) semi-urban and rural (girls) only. The estimated reliability coefficients by different methods ranged from 0.83 to 0.87 (V score), 0.76 to 1.79 (Q score) and 0.89 to 0.94 (total score). The standard error of measurement varied from three to five points. Intercorrelations of the four tests were also computed, which ranged from 0.44 to 0.69. Validity coefficients against the annual examination marks in Gujarati, mathematics and academic subjects only had the respective range from 0.29 to 0.68 (V score), from 0.30 to 0.75 (Q score) and from 0.42 to 0.75 (total score). Contingency coefficients against the teachers' rating were 0.57 (language teachers) and 0.51 (mathematics teachers). Validity coefficients with other allied psychological tests ranged from 0.59 to 0.77. All these coefficients were found to be statistically significant beyond 0.01 level of significance.

1072. SHARMA, B.B., *Experimental Verification of Various Methods of Examination in History at the Lower and Higher Stages,* Ph.D. Edu., Jammu U., 1977

The objectives of the study were: (i) to compare the reliability and validity of different types of examinations, namely, essay type, short answer, open book, objective type and viva voce, in history at lower and higher levels, (ii) to examine the reliability of examiners, (iii) to indicate the relative suitability of various forms of examination for the measurement of achievement in history, (iv) to appraise the significant differences between the scores by internal and external examiners, and (v) to familiarize teachers in general and examiners in particular with the recent trends and proposals of examination reforms.

The cluster sampling design was used to collect data from students and random sampling was used for

examiners. The sample consisted of 65 students, of whom 40 were from Class IX and 25 from the M.A. class. The sample of examiners included ten examiners of Class IX and ten of the M.A. class. For the collection of data, the assessment procedures used were essay type, short answer type, open book examination with essay-type questions, viva voce, objective type, questionnaire and observation. Analysis of the data was carried out using correlation techniques, factor analysis and three-way analysis of variance (3×3×3) in a factorial design at both the levels of the experiment.

The major findings of the study were: (i) The internal examiners discriminated better between examinees in essay, short answer, open book and viva voce examinations than the external examiners. (ii) There were wide differences in the marking standards of the examiners at both the levels in essay, short answer and open book examinations. (iii) Short answer tests yielded more reliable results at both the levels. (iv) The essay type worked well at the higher stage whereas the open book examinations did not work well at both the levels. (v) Grades reduced the error of subjectivity in the examiners and could be introduced at the postgraduate stage but the importance of marks and grades was considered almost equal at the high school stage. (vi) High factor loadings showed that grades could possibly be used as a scale for assessment. (vii) The factorial study indicated that history teaching includes all the eleven objectives studied in a lower or higher degree at both the stages but the emphasis on the objectives was different at different stages. (viii) Factor analysis showed that the second factor appeared polarized between intellectual versus emotional approach at the high school stage whereas the same appeared polarized between relaxed liberal versus radical approach at the postgraduate stage. (ix) The style of writing, length of answers and handwriting influenced the marking style of examiners at the high school stage whereas at the postgraduate stage only the style and handwriting influenced the marking style of examiners.

1073. SHARMA, H.R., *Construction, Administration and Standardization of an Achievement Test in General Knowledge (History and Geography) for Matric Students in Jammu Province*, Ph.D. Edu., Jammu U., 1981

The objectives of the study were: (i) to standardize an achievement test in general knowledge (history and geography) for Class X students of Jammu province, and (ii) to find out the differences in the average scores of boys and girls as well as between the mean scores of students belonging to rural and urban localities. A representative group of 1,600 students formed the sample.

The major findings of the study were: (i) The reliability coefficients of the general knowledge test, computed by applying three different methods, viz., split-half, K-R 21 and Flanagans' formula, were 0.96, 0.84 and 0.76, respectively. (ii) The validity of the achievement test was estimated to be 0.78. (iii) Statistically significant differences were established between the average scores of boys and girls, the difference was in favour of girls. (iv) The performance of boys and girls of the urban schools was superior to that of their counterparts studying in the rural schools.

1074. SHIVANANDA, D.S., *Construction and Standardization of a Reading Test in Kannada for Standards V, VI and VII*, Ph.D. Edu., Ban. U., 1981

The study had the objective of developing and standardizing a Reading Test in Kannada for pupils of Standards V, VI and VIII separately.

The test for each standard consisted of five subtests covering five aspects, namely, vocabulary, opposites, word discrimination, reading comprehension and sentence formation. The material selected for the tests for tryout was based on the textbooks and the vocabulary lists prepared by the Directorate of Textbooks with the technical collaboration of the NCERT. The preliminary forms of the tests were tried out on a representative sample of 100 students of each standard. After refining the items, they were again tried out on a representative sample of 188 pupils from each standard, drawn from nine schools covering both rural and urban areas. After item analysis, the final form of the test comprised thirty items for each subtest totalling 150 items in all for the whole test for each standard. For the purpose of developing norms, the test was administered to 2,036 pupils of Standard V, 2,010 pupils of Standard VI and 2,040 pupils of Standard VII. The sample included boys and girls, from rural and urban places studying in government and private schools. Percentile rank norms, t-scores and stanines were developed.

The validity coefficients ranged from 0.60 to 0.70 for Standard V, from 0.60 to 0.69 for Standard VI and from 0.50 to 0.56 for Standard VII. The test-retest reliability ranged from 0.582 to 0.920 for Standard V, from 0.584 to 0.885 for Standard VI, and from 0.602 to 0.937 for Stan-

dard VII. The split-half reliability ranged from 0.848 to 0.960 for Standard V, from 0.842 to 0.962 for Standard VI and from 0.877 to 0.964 for Standard VII. The intercorrelations for all the subtests for all the standards were also computed.

1075. SINGH, S., *A Study of Preparatory Science Examination Results of the M.S. University of Baroda for the Year 1974-75*, Ph.D. Edu., MSU, Baroda, 1978

The aims of the investigation were: (i) to study the distribution of marks for normality, (ii) to study the main sources of variations involved in examination marks, (iii) to study the relationship between marks under each of the procedures of assessment and prediction of examination marks from the previous examination marks, and (iv) to study the abilities underlying the examination.

The sample was drawn from the students who appeared in the preparatory science examination conducted for the year 1974-75. In all, 1,054 students who were studying at the university campus, were included in the sample. Data were obtained by collecting the marks of S.S.C. examination from the students' admission forms available at the office of the preparatory science unit. Preparatory Science (P.Sc.) marks were collected from the examination section of the university. The marks in all the papers were converted into percentage marks to bring them on the same scale. The marks were tested for normality by applying the chi-square test. Variations in the marking were studied by applying the t-test. The relationship and prediction were studied by using product moment correlations and step-wise multivariate regression analysis, respectively. The principal component method of factor analysis was applied for identifying the factors.

The major findings of the investigation were: (i) The distribution of marks, when studied separately in theory, tutorial and practical examinations, was non-normal. The distribution of preparatory science aggregate percentage marks was approximately normal. (ii) In theory as well as in tutorials, the standard of marking of different papers in different subjects differed significantly at 1 per cent level. (iii) The relationships between tutorial marks in diferent subjects, theory marks in different subjects and practical marks in different subjects were significant at 1 per cent level. (iv) The marks in the S.S.C. examination had got high predictive value for predicting the performance in the preparatory science examination. The S.S.C. aggregate percentage marks were the single best predictor for predicting the success in the preparatory science examination. (v) The underlying abilities measured by tutorial and theory examinations, separately, were the same. (vi) The abilities identified under the preparatory science examination indicated that the examination was comprehensive in nature and measured many dimensions of achievement abilities. (vii) In theory as well as in tutorial examinations, the common factors were general scholastic, scientific verbal and scientific numerical, whereas in practical examination, there was only one common factor of general ability. When the marks of all the three assessments were added, one more factor, namely, problem-solving, was found. The marks in the S.S.C. examination indicated an independent common factor named as memorization.

***1076.** ŚKARIAH, M.T., *Construction and Standardization of Oral Reading Comprehension Test in English for Pupils of Class VIII of Gujarat*, Ph.D. Edu., SPU, 1981

The aim of the study was to construct and standardize an oral reading comprehension test in English for pupils of Class VIII of Gujarat. The major objectives of the study were: (i) to measure the oral reading comprehension and oral reading errors of Class VIII pupils of Gujarat, (ii) to prepare different reading passages appropriate for Class VIII pupils for the purpose of oral reading with a view to measuring the oral reading comprehension, (iii) to establish various norms, calculate the reliability and determine the validity, and (iv) to undertake various concomitant studies pertaining to oral reading.

The reading material was selected after determining the readability of the text as well as the readability of the passages selected for the purpose of constructing the test. In the pilot form there were eighteen passages and eighty-six test items. On the basis of the pilot administration, discriminative values and difficulty values of the test items were computed. Keeping in view the readability of the passages and the discriminative value and the difficulty value of the test items eight passages were selected for inclusion in the final form of the test. The final form consisted of nine subtests including one for practice. The final form of the test was administered to 1,250 students selected by the stratified random method.

The main features and conclusions of the study were: (i) Separate percentile and standard score norms for boys and girls for reading comprehension and error

scores were determined. (ii) The reliability of the test, established by test-retest, split-half and K-R formula, ranged between 0.82 and 0.97. (iii) The concurrent and predictive validities had been established by correlating the teachers' opinions and the marks obtained in the examination with the scores on the oral reading test. (iv) There were significant sex differences and area differences. Girls were better in comprehension and made fewer errors in oral reading.

1077. SIERT, *Analysis of Middle Standard Examination Results — 1964-1968,* Karnal, 1970

The specific aims of the study were: (i) to investigate the comparative trends of overall and subject-wise pass percentage of Class VIII examination results, (ii) to gauge the relationship between the marks obtained in English and Hindi and between mathematics and general science, and (iii) to find out the causes of failures. Three questionnaires were administered to all the schools in the State. In addition to this D.E.Os also sent the results of 8,603 students.

The major findings of the study were: (i) The pass percentages of the results over the five years under study were fairly satisfactory. However, the performance in general was poor. The average of the marks obtained by 875 students in the year 1968 was only 36.1, which was very near the minimum pass marks—33. (ii) There was a constant downward trend in the pass percentage in social studies and these had considerably fallen in the year 1967. (iii) The result of general science was also low. (iv) The level of performance of Mohindergarh district students had been the lowest all through the period 1964-68. (v) The pass percentage of Gurgaon district had a constant downward trend in pupils' level of performance. It varied from 92.8 in 1964 to 73.8 in 1968. (vi) There was a high incidence of marginal failures and marginal passes. (vii) The percentage of the students securing about 60 per cent marks in Hindi, general science and social studies ranged between one to four, which was very low. This low percentage of bright students indicated that due attention was not paid to the bright students. Probably more emphasis was paid to quantity rather than quality. (viii) There was a positive and significant relationship between Hindi and English and between mathematics and general science.

1078. THAKORE, B., *Construction of Diagnostic Tests and Preparing Remedial Material as well as Testing Its Effectiveness on Fractions and Decimal Fractions for the Students of Grade V of Gujarati-medium Schools in Greater Bombay,* Ph.D. Edu., SNDT, 1980

The major aim was construction of diagnostic tests followed by preparation of remedial material and testing its effectiveness. The tests and the follow-up remedial material were prepared for students of Class V on two topics, fractions and decimal fractions.

In all, ten diagnostic tests were prepared. Tests were administered to the students of eleven schools of Bombay, selected randomly. They were analysed for errors. The remedial teaching material was prepared and was used with the same students. The posttest, which was also like the pretest, was given to find out the effectiveness of the remedial material prepared.

The major findings were: (i) The students of Class V did not have clear concept of fractions. It was difficult for them to convert one type of fractions into another. Simplification was found to be a difficult process for them. They could not easily find out the smallest and the biggest fraction. Many simple principles, like zero multiplied by any fraction is zero, division is the inverse of multiplication, and the practical use of fractions, were not clear to them. (ii) As for decimal fractions, the students did not understand the place values of respective figures in decimal fractions. (iii) They did not understand addition, subtraction, multiplication and division of decimal fractions.

The investigator prepared a list of suggestions for improving and correcting the above-mentioned points in respect of the learning of fractions and decimal fractions.

***1079.** THAKORE, B.M., *The Development of Diagnostic Tests on Decimals and Fractions for Students of Class V of the Gujarati-medium Schools of Greater Bombay and to Study the Effectiveness of Instructional Material Prepared for Them,* Ph.D. Edu., SNDT, 1980

The major aims of the investigation were: (i) to analyse the errors committed by students of Class V of the Gujarati-medium schools of Greater Bombay, (ii) to develop diagnostic tests, and (iii) to develop remedial programmes and try out their effectiveness.

The investigator selected fractions and decimals and analysed the errors committed by students in dealing with the examples on these topics. The standard

methodology of constructing a diagnostic test was used by the investigator. The various sub-units on which test items were developed were understanding the meaning, additions and subtractions, problems based on additions and subtractions, multiplications and divisions. A pre-tryout of the test was undertaken on sixty students from three schools of Bombay. As a result of pre-tryout, the test items were revised and edited. The revised tests were tried out on a sample of 218 students selected from four different schools of Bombay. Based on the difficulty levels of the items, certain items were rejected and others retained. The final tryout of the test was undertaken on a sample of 581 students of eleven schools. Descriptive statistics and critical ratio were used for the analysis of the results. Anaylsis of the errors committed by the students was undertaken. Based on the analysis of errors, remedial assignments were prepared.

1080. UPADHYAYA, I.J., *Standardization of the Proficiency Test in First Language for Primary School Teachers of Gujarat State,* Ph.D. Edu., SPU, 1979

The main objective of the study was to construct and standardize a language proficiency test for the primary school teachers of Gujarat.

In order to construct a valid test of language proficiency, the terms were properly defined by reviewing a few definitions given by the experts in the field. The components and areas for the Language Proficiency Tests with their behaviour components were specified. After the pre-pilot and pilot administration of the test items, one hundred and fifty items were selected. The test items were properly clustered into subtests in major areas like vocabulary, written expression and reading comprehension. The test was administered to 4,597 primary school teachers.

The reliability was established by various techniques. The reliability coefficient was 0.82 by test-retest, 0.98 by analysis of variance, 0.90 by split-half and 0.94 by formula K-R 20. The content, concurrent, predictive and factorial validities were established. Separate norms for various qualifications were shown in the test manual.

1081. VENKUBAI, J., *Internal Assessment: Use and Misuse,* Directorate of Higher Education, Hyderabad, 1965

The study was undertaken to find out up to what extent the 25 per cent weightage of class record given at the public examination was responsible for the high percentage in mathematics and science.

The sample consisted of twenty-five schools from Hyderabad and Secunderabad. Since the investigation was confined to the city schools, all the schools were taken up for the study without any omission. A proforma was prepared to gather certain details from the selected schools. The proforma was administered to collect information on the performance of the pupils in Class XI during 1962-63 and in Class XII during 1963-64 before the pupils took their examinations. All the schools were visited during the months of July to September 1964. The heads of institutions and experienced teachers were interviewed. Advantages and disadvantages of having internal assessment were discussed with them. School records relating to the assessment work were scrutinized. These included cumulative marks registers, answer scripts of monthly and terminal tests, question papers, pupils' notebooks and practical records.

The study yielded the following findings: (i) Out of the twenty-five schools, nine schools resorted to inflation of marks in the internal assessment in mathematics. Out of these, only four schools got the advantage of higher average or increased number of passes in the subject. (ii) In certain schools the heads of the institutions devised improved methods in the maintenance of records. The form of internal assessment varied from school to school and from subject to subject. (iii) In physical sciences, seven schools showed inflation in the assessment of classwork. (iv) In biological sciences, eight schools, out of the twenty-five showed higher percentage of internal assessment than the average performance of the candidates at the public examination. (v) A major part of the internal assessment was based on written work alone — slip tests, terminal examinations or assignments. (vi) Out of the twenty-five heads of the institutions, only five were positively against the continuance of the internal assessment. Among others three were vague and indecisive. The remaining seventeen were positively in favour of having internal assessment. (vii) There was no uniformity in the mode of practical examinations conducted. It varied from school to school according to facilities and equipment available. (viii) In biological sciences, eight schools out of the twenty-five showed higher percentage in the internal assessment than the performance at the public examination. (ix) The internal assessment was not responsible for higher percentage at the public examination.

1082. VERMA, L.K., *A Critical Appraisal of Some In-*

novations for the Improvement of Examinations, Ph.D. Edu., Kur. U., 1981

The major objective of the study was to make experimental validations of some innovations implemented in the examination system of universities. These innovations included marking and grading, internal and external assessment, improvement of the essay-type question paper, open book vs. traditional examination, internal vs. external choice in questions, spot evaluation and revaluation.

Different procedures and sampling processes were followed to validate different innovations of the study. For marking vs. grading, fifty answer scripts of B.A. (history) were got evaluated on a five-point scale, seven-point scale and numerical scale of 101 points by five examiners. For internal and external assessment, internal and external assessment scores were collected for honours course of Utkal University, M.Ed. course of Jammu University, M.A. course of Sardar Patel and Bombay Universities, and M.B.A. course of Kurukshetra University. For the improvement of essay-type question paper, the question paper already set in the annual examination was reframed in simple, clear and precise language. It was administered to two groups of randomly selected fifty students. The answer scripts of the students of the two groups were evaluated by five examiners. In the case of open book vs. traditional examination, an objective-type test in chemistry was administered to pre-university students. They were also administered the Sinha Scientific Aptitude Test. The scores of the open book examination and the university examination were correlated separately with the criterion test and the significance of difference between the two correlations was tested. For internal vs. external choice in questions, a question paper with ten short answer type questions in two forms was administered to the students. In one form the students had choice for all the ten questions and in another form the questions were in pairs. The answer books of the students were evaluated separately by three examiners. For spot evaluation, a self-prepared attitude scale was administered to 146 approved examiners of the universities in arts and science courses at the undergraduate level and the postgraduate level. These examiners had experience of on-the-spot evaluation. For revaluation, rules and regulations of eight universities of northern India were compared.

The findings of the study were: (i) Though there was no significant difference between the average assessment on five and seven-point scale, better discrimination was possible through the latter. (ii) The internal assessment scores highly correlated with the external assessment scores of the only two universities where the number of students was small. (iii) The inter-examiner reliability of the scores awarded by different examiners improved with the improvement of the question paper. (iv) The correlation coefficient between the scientific aptitude test and the essay-type test was low. (v) The scores of the students in internal choice questions had a significantly greater content validity than that of the scores in external choice questions. (vi) The arts and science examiners at the undergraduate level or the postgraduate level did not differ from each other on the different items of the spot evaluation attitude scale. (vii) There was no uniformity in the rules and regulations regarding revaluation in different universities; further, the different rules in a particular university might affect the same case differently.