13

Educational Evaluation and Examinations

A Trend Report

B. K. PASSI
D. N. SANSANWAL

According to the Education Commission (1964-66), examinations and evaluation constitute one of those areas in education about which one can say that the problem is known, and its significance is realised... A number of institutions in the country have been undertaking sustained studies in the area of examinations and evaluation during the last twenty years. Buch and Lele in the M. S. University of Baroda, Gayen at the IIT, 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 UGC has supported willing universities in setting up examination research and reform units. It has given substantial grants and support to twelve universities for improving examination in higher education. In the document of UGC (1973) "Examination Reform — A Plan of Action", it is expressed that examinations have dominated educational process, and external examinations have encouraged selective study and cramming. Examination marks have lacked reliability and validity, and unfair means in the examinations have increased tremendously. The crippling effect of external examination on the quality of instruction has compelled various agencies to bring 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. The NCERT and some boards of secondary education have also geared their activities to conduct researches and implement various innovative programmes in the examination system.

A substantial amount of work has been done with the grants received from the NCERT and the UGC in a number of universities. These studies have been compiled and/or reviewed by Long and Mehta (1966), Pareek and Kumar (1966), Pareek and Sood (1971), Dave (1968), and Buch (1972). 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 in all. Of these, one hundred were classified under intelligence, ninetysix under achievement, sixty under aptitude, fifteen under interest, fortyfive under personality and ten under miscellaneous category. Out of the ninetysix tests under achievement, thirtyfive 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, twentyeight on other subjects and ninetythree 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 split of these studies is as follows : examination — seventythree, general achievement — fortyfour, language — fortyfive, mathematics — fortytwo, social studies — fortytwo, science — twentythree, others subjects — twentyfour, and scholastic backwardness — thirtythree.

Dave (1968) in the Third Indian Yearbook 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. The category-wise split of the source material is as follows : M.Ed. dissertations about eightyeight percent, Ph.D. studies about two percent, published research papers about two percent and projects about eight percent. The above data indicate that it is exclusively dominated by research studies conducted at the M.Ed. level. Dave (1968) reviewed only ten Ph.D. studies, while at that time, i.e., by 1968, probably nineteen Ph.D. studies in education with seven more Ph.D. studies from other disciplines but related to the field of evaluation were available in the country. While categorising studies reviewed by Dave (1968) according to the structure of this review, it is found that achievement tests, diagnostic tests, examinations, factors affecting achievement, prediction-admission-promotion and failures cover about fiftyone, six, eighteen, nineteen, three and three percent respectively. Based upon the data from two types of information of Dave's review, viz., (i) the nature of source material and (ii) the nature of studies, as classified above, it can be inferred that most of the studies in the area of evaluation in Indian educational research are in the field of achievement test construction at the M.Ed. level. It is for the future researchers and research consumers to see how to intergrate and utilise this vast source material.

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.

The present trend report is based upon 127 studies available in the area of examinations and achievement testing both at Ph.D. and at project levels. These studies have been classified into six different areas which can be distinguished from each other in terms of emphasis only. The six areas are : (i) Achievement Tests, (ii) Diagnostic Tests, (iii) Factors Affecting Achievement, (iv) Examinations, (v) Prediction-Admission-Promotion, and (vi) Failures. The subsequent discussion in this trend report will follow this sequence.

NATURE OF STUDIES

The present review is based upon Ph.D. work done in Indian universities and research projects completed upto April 1976.

A detailed picture of the six subareas of the problems of research in Educational Evaluation and Examinations and the developmental trends is given in Table 1.

Table 1 includes data indicating periodwise growth of research work in the subareas of Educational Evaluation and Examinations. It may be noticed here that upto 1954, very little research work has been done and it is restricted to two subareas, viz., Factors Affecting Achievement and Prediction-Admission-Promotion. In the years 1960 to 1974 there is a heavy concentration (about two-thirds) in the subarea of Achievement Tests. No research in the area of Diagnostic Tests has been done upto 1964.

TABLE 1

AREAWISE AND PERIODWISE DISTRIBUTION OF STUDIES

	Upto 1944	1945-49	1950-54	1955-59	1960-64	1965-69	1970-74	1975+	Total
Achievement Test	—	—	—	2	10	15	15	3	45
Diagnostic Tests	—	—	—	—	—	5	4	—	9
Examinations	—	—	—	2	11	10	10	—	33
Factors Affecting Achievement	1	—	2	5	2	4	3	—	17
Prediction-Admission-Promotion	—	—	2	2	6	3	2	1	16
Failures	—	—	—	2	4	1	—	—	7
Total	1	—	4	13	33	38	34	4	127

The studies related to Failures have been 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 at the Ph.D. level before independence. After independence there is a gradual increase in Ph.D. and institutional research in this area. During the period 1965-69, thirtyeight studies were completed. Another interesting feature is that out of 127 studies during the period 1944-1975, 105 studies were completed during the period 1960-1974. Achievement Tests claim 45 studies and Examinations claim 33 studies. The researchers, it seems, are not attracted to study the problem of Failures where only seven studies are cited. Even these seven studies are institutional projects rather than Ph.D. research.

METHODOLOGY

Sample :

Different investigators aiming at standardising tests have used samples of different sizes and natures. About two-thirds of these tests have been standardised on samples of less than three thousand students and there are only a few such test standardisation studies which have used samples of more than seven thousand students. The samples from 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 randomisation, multistage randomisation, stratified sampling, random stratified, clustered, etc., depending upon the nature of the related population and the purpose of the study.

Statistical Designs :

From the methodological point of view, the break-up of 127 studies is as follows : (i) descriptive and correlational—106, (ii) factor analytical—nine, (iii) regression—nine, (iv) experimental—two, and (v) regression and factor analytical—one. The two experimental studies have used simple pretest and post-test 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 S.S.C. Examination marks and criterion like college grades.

ACHIEVEMENT TESTS

The review of studies in 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 :

In the domain of scholastic achievement, efforts have been made by Lele, Parikh, Palkar and Parikh (1964b), Liddle (1965), Jha (1974) and Sharma (1975). Lele et al. (1964b) prepared a Scholastic Aptitude Test for admission to preparatory science courses. The test consisted of three subtests on English, numerical ability and abstract reasoning. Liddle (1965) standardised 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 ranged from 0.83 to 0.89. The concurrent validity coefficients against scholastic achievement in terms of total scores, ranged from 0.464 to 0.766. Jha (1974) developed a battery of tests for measuring basic skills related to arithmetic, geometry, and languages. Sharma (1975) developed a biological science aptitude test for school students.

Languages :

In the case of achievement tests for English, there have been only eight tests for grades VI to XI. Aram, Rangaswamy and Feroze (1957) have standardised an achievement test in English for middle stage students of Coimbatore district of Tamil Nadu. The test included subtests related to language usage, spellings, punctuation, capitalisation, reading, comprehension and vocabulary. Buch, Patel and Kotwal (1960) standardised 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, Automatization 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.

Patel (1971) standardised a Silent Reading Comprehension Test in English for S.S.C. pupils of Gujarat. The test contents were related to ability to note significant details, to select appropriate meaning of the word in content, to read maps and tables, to follow sequence of events, to draw generalisation, 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 test shared loadings with three factors, viz., (i) Ability to Grasp, (ii) Word Meaning, and (iii) Perceptual Speed Factor.

Shukla and Tutoo (1959), CIE (1962), Gujarat Research Society (1963), Jha, Altekar, Jha, Gajare and Indapurkar (1964), Sharma (1967), Deshpande (1972), and Gaur (1973) constructed achievement tests in Hindi. These studies provided tests for grade V to S.S.C. level. It is to be noted that achievement test construction in Hindi is localised in Delhi, Haryana, 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 City. Jha et al. (1964) developed achievement tests in Hindi for students of grades V to VIII. The subtests included were vocabulary, comprehension, recognition of tenses, spellings, sentence structure, 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 students of grades VIII to X, whereas Deshpande (1972) made similar efforts for higher secondary school students of Maharashtra. Gaur (1973) standardised achievement tests in six different aspects of Hindi, viz., spelling,

vocabulary, applied grammar, formal grammar, idioms and proverbs, and comprehension. The test was standardised for students of Matriculation level of Haryana. The split-half reliability coefficients for the tests varied from 0.83 to 0.92.

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), Gujarat Research Society (1963), Bhatt (1971), Krishnamurti (1971), Maniar (1973), Pandya (1973), Parekh (1973), Desai (1974), Gohil (1974), and Modi (1975). Achievement tests in Gujarati are available for all grades from V to pre-university and for three to five years age group children. Buch et al. (1960) constructed achievement tests in Gujarati for 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 children of Bombay of age range thirteen to seventeen. The Gujarat Research Society (1963) developed achievement tests in Gujarati for children of Bombay of standards V to VII. Bhatt (1971) developed an Achievement Test in Gujarati for students of grade VIII and prepared the norms for Gujarat State. 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 school examination marks in Gujarati was found to be 0.86. Pandya (1973) standardised a Language Ability Test in Gujarati for 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 S.S.C. Examination was 0.75. Desai (1974) standardised Language Development Tests for Gujarati children of age group three to five. The test includes items related to comprehension, sound discrimination, articulation, and oral expression. Bhagatwala (1960) standardised a Silent Reading Test in Gujarati for grades VIII to XI. The test intended to measure speed of reading, word meaning and comprehension. Krishnamurti (1971) at Madras also developed Reading Readiness Test for school children. Maniar (1973) developed a Silent Reading Test in Gujarati for 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) standardised 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. Gohil (1974) standardised 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 standardisation 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.

There are two other studies related to achievement test consruction in two other regional languages —Oriya and Marathi. Dash (1967) standardised an achievement test in Oriya for grade VII students, whereas Deshpande (1972) developed an objective assessment tool in Marathi for students appearing for the secondary education examination in Maharashtra.

To sum up, the achievement test construction in languages at national and regional levels needs more extensive and concentrated efforts. Except in three regional languages, namely, Gujarati, Oriya and Marathi, no systematic and sustained attempts have been made to standardise tests in other regional languages at the Ph.D. level or at the institutional level. It is interesting to see that the studies in the areas of reading speed, reading comprehension, and listening comprehension are restricted to Gujarati language. All studies have been localised in Gujarat except one study by Maniar (1961). Research studies at the M.Ed. level aiming at measuring achievement in regional languages should be coordinated to take care that each regional language has achievement tests for all grade levels. In the case of Hindi, a national project should be undertaken on the lines of the study 'All India Survey of Achievement in Mathematics' conducted by Kulkarni, Mohanlal and Naidu (1970).

Social Sciences :

Achievement tests in the area of social sciences were constructed by Aram et al. (1957), Shukla and Tutoo (1959), Buch et al. (1960), Gujarat Research Society (1963), Saraf (1964), SIE, Kerala (1965), Dash (1967), Srivastava (1967), Misra (1968), Vanajakshi (1970), Misra (1970) and Deshpande (1972). They developed tests in social studies, history, geography and civics only. The social studies tests were

constructed for students of grades IV to VIII, history for grades V to XI, geography for grades V to VIII, X and XI, and civics for grade X only.

Aram et al. (1957) developed an achievement test in social sciences for middle school students of Coimbatore district in Tamil Nadu. SIE, Kerala (1965) standardised an achievement test in social studies for students of standard VIII of Kerala. The test had a split-half reliability of 0.72. Similarly, Vanajakshi (1970) developed a test for 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 students of grade VII, an achievement test in social studies was standardised by Dash (1967) for the students of Orissa.

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. Gujarat Research Society (1963) standardised tests in history for students of grades V to VII. Misra (1968) standardised an achievement test in history for high school students of Uttar Pradesh. The contents of the test covered 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 Delhi students of grade XI, an achievement test in Indian history was standardised 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 coefficient against the criteria of school board examination marks, teacher ratings, etc., ranged between 0.28 and 0.68. An attempt in this direction was also made by Misra (1970) for students of grade X in Assam.

For geography at school stage, Buch et al. (1960), Misra (1970) and Deshpande (1972) developed achievement tests for grades VIII to X of Gujarat, for grade X of Assam, and for grade XI of Maharashtra, respectively. Achievement tests in geography were also standardised for grade V to VII by Gujarat Research Society (1963). For achievement tests in civics at school stage, only one attempt has been made by Srivastava (1967) for high school students of Uttar Pradesh.

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 standarised at all.

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 the states of 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), 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 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 into corresponding coefficients of 0.48 and 0.62. SIE, Kerala (1965) standardised an achievement test in general mathematics for 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 standardising an achievement test battery for 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 standardised 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 high school students of Assam state. Bhatt (1971) standardised an achievement test in mathematics for grade VIII students of Gujarat. The test-retest, split-half, K—R formula 20 and Rulon's methods resulted into 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 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 different grades from III to X for the states of Maharashtra, Gujarat, Mysore and West Bengal. Chickemane (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 standardised a group test of problem solving ability in arithmetic. Dave (1958) constructed achievement tests in arithmetic for students of different grades of Bombay. The test-retest reliability coefficient was 0.985 and validity coefficient against 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. Gujarat Research Society (1963) and Pendharkar (1965) standardised tests for students of grades V, VI and VII of Greater Bombay. Jha (1974) developed tests related to arithmetic concepts and certain concepts of geometry for measuring basic skills of students of classes 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.

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 achievement scores as the external criterion of validation. There is a need for establishing other types of validity coefficients.

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), Gujarat Research Society (1963), SIE, Kerala (1965), Dash (1967), Sheth (1967), Rup Prakash (1968), Vanajakshi (1970), Bhatt (1971), and Hira Devi (1973). In the case of general science, tests are available for the 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 middle school students of Coimbatore district. 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, 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 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) standardised achievement tests in general science for the students of grades V, VI, and VII in Uttar Pradesh and Bombay respe-

ctively. The reliability coefficients calculated by split-half, Rulon's and K—R formula 21 methods ranged from 0.88 to 0.99 for all the different grades. Concurrent validity coefficients were worked out against teachers' ratings and annual examination marks. Vanajakshi (1970) standardised an achievement test in general science for students of grade VII of Andhra Pradesh. The test had 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 test-retest, split-half, K—R formulae 20 and 21, and Rulon's methods ranged from 0.92 to 0.94.

For the school subjects — physics and chemistry, Bountra (1970), Gupta (1974), and Tewari (1975) standardised the tests for high school and college students by employing samples from Uttar Pradesh and Haryana. Bountra (1970) developed the 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 school achievement were 0.60 and 0.61 respectively. Gupta (1974) developed six different types of tests in physics and chemistry for high school/pre-university students. The split-half reliability coefficients ranged from 0.78 to 0.95. Tewari (1975) standardised attainment tests in physics, chemistry, botany, and zoology for high school students of Uttar Pradesh. The factorial validity of the tests was worked out.

Kapoor (1968) and Garg (1969) standardised achievement tests in home science for high and higher secondary students of Uttar Pradesh respectively. The validity coefficients against 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 a 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 entrance tests which have been developed for the purposes of admission and classification. 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 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 a 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 students of multipurpose and higher secondary schools of Mysore state. 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 schools where RSSB was used for streamlining the students, the success ratio of schools ranged from 0.96 to 0.73. Exploring the area of physical education, Shukla (1957) developed a physical education test. He introduced the concept of 'Physical Age' and 'Physical Quotient' (P.Q.). Pillai et al. (1967) made efforts to assess handwriting by the use of the newly standardised Kerala University Handwriting Scale. De (1965) canstructed an achievement test of 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.

Although the area of achievement test construction has attracted a great deal of attention from Ph.D. scholars and other investigators, it can be safely remarked that there is a lack of planning and coordination in the work done. A systematic and continuous work over a period of time should result into workable tools for different school subjects, grades, regions, etc. There should be some centrally organised 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 are only nine research studies out of which only three are Ph.D. studies. Bombay University, Kerala University, Kurukshetra University, Lucknow University, Bureau of Psychology, Allahabad and the SIE, Gujarat are the institutions which have paid attention to this area. It is only very recently, that is, in 1966, when the first attempt was made by Mehta (1966) to undertake research work at Ph.D. level in the area of Diagnostic Tests. These projects are only for languages — Hindi

and Gujarati, and mathematics — arithmetic, algebra and geometry. There is none for social studies or physical sciences.

Languages :

There are four studies employing the diagnostic tests in languages. Sinha (1971) constructed a diagnostic test in Hindi for 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 standardised by I.T. College, Lucknow (1971). The test was for 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 concurrent validity coefficient was 0.69. Gomathy Ammal (1972) conducted a study to locate specific difficulties of 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.

Mathematics :

In arithmetic, Mehta (1966) constructed a diagnostic test related to the use of four fundamental rules of arithmetic. SIE, Gujarat (1969b) also did exactly a similar work. Both these studies were followed by remedial teaching which resulted in significant improvement. Sharma (1969) constructed a diagnostic test in algebra for 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 students who failed in grade X. Ashar (1972) standardised a diagnostic test in basic algebraic skills for Gujarati medium students of grades VIII, IX and X from Greater Bombay. The median reliability coefficient was 0.90. The concurrent validity coefficient against the scores in mathematics at annual examination at VIII, IX and X grades were 0.94, 0.91 and 0.98 respectively. It was found that pupils committed errors due to 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, the Diagnostic Test in the Skill of Using Geometrical Instruments was developed by the SIE, Gujarat.

In some of the studies, the procedure in the diag-

nostic test construction has been followed in the same way as that in the case of achievement test construction. The purpose, content, format and use of diagnostic tests are different from those of achievement tests. Consequently, the methodology to be followed in the case of diagnostic test construction should be different.

EXAMINATIONS

Out of the 127 studies included in the area of Educational Evaluation and Examinations, as many as thirtythree studies have been classified in the subarea of Examinations. These studies are related to general aspects of research on examinations, like public examinations, internal assessment, rate of chance in traditional examinations, etc. Except for the five Ph.D. studies by Rao (1968), Misra (1970), Deshpande (1972), Shah (1972), and Tluanga (1974), all the other studies are institutional projects conducted at Kharagpur, Gauhati, Baroda, Allahabad, Poona, New Delhi, Anand, Bombay, Annamalai, Nagpur and Saugar. The institutional projects have been conducted by Bokil (1956-63), Shukla (1959), Lele et al. (1962-1963), Taylor (1963), Taylor and Tluanga (1963), Taylor (1964a, 1964b), Raina (1964), Dave and Patel (1966), Sharma (1966), Kamat (1968), Jhaveri and Patel (1968), Patel (1968), Rao (1968), Gayen et al. (1961-1970), Harper (1962-1970), Nath (1972, 1974), Deo (1974) and Tluanga (1974).

A significant group of studies conducted by Gayen et al. (1961-70) at 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 appears, could be made in the first area of the studies of this report, i.e., Achievement Tests. But no where have these authors constructed and standardised 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) grouping of items and balancing of alternatives; (iii) difficulty values and discriminating powers of the question items; (iv) relationship of marks in a particular item with the total marks; (v) comparative study of 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 because 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.

Researchers like Harper (1962), Jhaveri and Patel (1968), Misra (1970), Taylor (1963, 1964a, 1964b), Nath (1972), Deo (1974), and Tluanga (1974) have 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 as compared to traditional examinations. Harper (1970) conducted his 'ninety marking ten' experiment and found that inter-examiner reliability was low. The low inter-scorer reliability was attributed to the essay type of examinations, types of questions, inadequate instructions and training of examiners. Jhaveri and Patel (1968) found that intra-examiner reliability increased more significantly in the case of well defined essays than traditional essays. Misra (1970) found inter-examiner reliability coefficient of 0.70 for essay. Taylor (1963) made an extreme comment that an examiner's mark had neither the sanctity nor the precision which was usually attached to it. Different examiners showed a large variation in the mean and standard deviation of their evaluation. Taylor (1964a), in his study, 'An Examination of Examiners', emphasised 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 passed the examination. He said that it was much more likely than before that these candidates were the ones who ought to have

passed. Nath (1972) also found that scaling reduced the dispersion of marks distribution, thus bringing down the variations between the examiners. In this context two studies of Deo (1974) and Tluanga (1974) are interesting to note. Deo (1974) in her study, namely, '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. 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. 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.

Very few studies by Lele et al. (1962b, 1963a), Chauhan (1967), Rao (1968) and Malhotra (1972) have been completed in relation to question paper and their nature. Chauhan (1967) in his study of university examination found that, of the students who fail, a large proportion fails not because of not knowing the subject matter but because of some external factors like defective question papers, carelessness of examiners in evaluating answer books and evaluation by incapable examiners. In a similar context, Malhotra (1972), while studying the effectiveness of question papers of matriculate 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 the question papers and found that only one-fifth of the total questions were good, whereas the rest were poor discriminators. Rao (1968) investigated into the system of tests and examinations at standard XI. He found that in many cases the question papers were not properly balanced as far as the difficulty value of the items was concerned. The question papers even failed to discriminate pupils of high from low ability. Further more, Lele et al. (1963a) found that in essay types 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. Hill (1964) found 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.

The difficulties regarding the assessment through the present system of examinations are manifold. Dave and Patel (1966) found that there was a substantial variation among the pass percentage of different high and higher secondary school examinations in a given year. They found a substantial increase in variation among pass percentage over a five year period from 1960 to 1964. Levels of achievement scores have different meanings depending upon the objectives of educational achievement as well as school subjects. 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. Shukla (1959) found that 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.

Studies by Raina (1964), Kamat (1968), Deshpande (1972) and Shah (1972) tried to find out the relationship between external and internal assessments and also their reliability. Shah (1972) surveyed the pattern of internal-external assessment 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 it was to increase. 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 the values of correlation coefficients between the external and internal assessment were greater in language subjects as compared to other subjects. 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 for second year engineering mathematics and —0.38 (not significant) for M.A./M.Sc. mathematics students.

The seminar held in 1971 on 'Examinations in Higher Education' suggested to concentrate upon issues related to redesigning of question papers, use of marks, grades and even percentiles as bases of awards, mass copying in the examinations, use of mechanical aids, strengthening examination wings, and introducing semester system of examination. Finally the seminar called upon the UGC and the Inter-University Board to organise a well staffed unit at the national level for research into problems of examinations. This unit should have its extension centres in other universities and colleges.

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 using of standard error of 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 future.

FACTORS AFFECTING ACHIEVEMENT

In this area seventeen studies have been included. They cover correlates of achievement and factorial structure of achievement domain.

Correlates of Achievement :

There have been studies investigating into the correlates of achievement, such as, 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 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. Misra (1970) said that out of ninety examiners scoring the same ten answer sheets, one of the examiners passed only ten percent of the students, whereas another examiner passed ninety percent 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. Studying the testee variables like age, sex, caste and locality of students in relation to achieve-

ment, different investigators like Bokil (1956a), Dave (1958) and Sharma (1967) made studies in Maharashtra and Gujarat. Bokil (1956a) found that the pass percentage of girls in 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 upto 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. Bokil (1956a) found that in the case of S.S.C. candidates, the pass percentage decreased with the increase in age to the extent that at sixteen, the pass percentage was seventyseven and at twenty, the percentage was twentyseven. Sharma (1967) found in the context of caste of students that with higher levels of caste there was better achievement. Bokil (1959a) found that number of languages learnt affected achievement.

Investigators like Bokil (1956b) and Sharma (1967) found that the variable 'medium of instruction', influenced achievement significantly. But SIE, Maharashtra (1971) found no significant difference in the achievement of students with English and non-English as their media of instruction. Bokil (1956a) found that the pass percentage of students appearing for S.S.C. examinations held in March and October differed significantly. Bokil (1959b) also found that pass percentage was highest in schools of medium size. Nath (1974) found that achievement of pre-university students was related to student-teacher ratio. Sharma (1967) found that 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 clearcut superiority of one type of school over the other. Sharma (1967), Bokil (1956b), Rao and Arunajatai (1971) found that 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 Panjab University. The above studies show that testing conditions and nature of test influence the achievement of students.

Factorial Studies in Achievement :

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. Factor analysis as a technique has been used by Chickermane (1943), Gokhale (1954), Mehrotra (1954), Lele, Parikh, Palkar and Parikh (1964a), Pendharkar (1965), Dash (1967), Sinha (1967), Kamat

(1968), Deshpande (1972), Gupta (1974) and Tewari (1975). 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 factorising 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.E. and P.Sc.E. 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., Automatisation, Comprehension-Vocabulary, Pronunciation and a fourth factor having loadings on grammatical aspects of English achievement. As regards the study on 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 assessment 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.

More efforts have to be made in this direction by covering large representative samples, using better factorisation methods and covering different facets of achievement in terms of educational objectives. Factor structure of achievement under convergent and divergent test contents and formats is yet to be explored.

PREDICTION-ADMISSION-PROMOTION STUDIES

How effective are the predictions? There can be three approaches to handle this question, 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 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 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 the Jalota's General Mental Ability Test was a good predictor for general science (r = 0.80). The Nafde's Nonverbal Intelligence Test was found to be an effective predictor for mathematics. He also used other tests like the 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 he 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 a 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 intermediate level. The multiple R was 0.49.

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 subject choice at degree level by those who secured less than forty-five percent marks should be made with caution. Higher secondary examination marks in languages and mathematics were useful for selection of courses and prediction purposes. Dharm Bir and Srivastava (1965) tried to find 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, English, etc. Buch (1963) found that S.S.C.E. composite scores of physics, chemistry, and mathematics were better predictors than the individual set of scores while predicting success in the next higher examinations. Hiriyanniah (1963), while finding the prognostic value of pre-university examinations, found that 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 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 unscaled marks turned out to be as good predictors as scaled marks.

Lele and Bhagatwala (1954) developed and used the University Entrance Test for admitting students to the Faculty of Arts and Faculty of Science. The test had a maximum predictive validity coefficient of 0.74 for arts students. If the nature of students and the nature of predictor and criterion variables remain stable, prediction studies will be more useful, but unluckily, 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 should be designed and undertaken.

FAILURES

Failure in examinations is a serious problem when viewed in terms of wastage and stagnation. Education Commission (1964-66) has reported 'At present, out of every 100 children who enter Class I, only about half complete Class IV and only 34 complete Class VII'. There can be several reasons for this — economic, social and educational. In this review, there are seven studies which are directly related to the phenomenon of failures. Bokil (1956c, 1958a and 1963) conducted three studies by drawing the samples from the S.S.C.E. Board of Maharashtra. The studies by DEPSE (1964), NCERT (1964), GCPI, Allahabad (1964) and NCERT (1965) are the other four institutional studies related to the area of Failures in Examinations. Bokil (1956c) found that 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 twenty percent of the total marks. For 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 Engilsh, Bokil (1963) found that one-third of the failures were by a margin of less than ten percent marks; overall average score in English was thirty-one, which was less than the pass marks of thirtyfive. Bokil (1958a) while calculating average percentages of failures in English, found that there were significant differences from group to group and from year to year. DEPSE (1964) found that teachers' qualifications and background, teaching methods, working conditions and location of schools, transfers, buildings, equipment, clerical work done by teachers, pupils' previous attainment, pupils' attendance, media of instruction, examinations, etc., were factors which were related to pass — fail percentage of schools. It was also found that the largest failures at 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, Allahabad (1964) also found favourable results for regular candidates. It was found that girls failed less than boys. Performance in English, mathematics, science and civics was responsible for higher incidence of failure. The GCPI found that many social, economic, and educational factors were responsible for failure in schools. The NCERT (1965) conducted a sample study of failures in boards of secondary examinations taking the sample from 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.

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 characterised as inadequate, not well integrated and, therefore, at moments not goal directed.

The researches appear inadequate when the importance of the field and the existence of gaps are considered. The gaps are conspicuous by the ab-

(1968), Deshpande (1972), Gupta (1974) and Tewari (1975). 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 factorising 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.E. and P.Sc.E. 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., Automatisation, Comprehension-Vocabulary, Pronunciation and a fourth factor having loadings on grammatical aspects of English achievement. As regards the study on 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 assessment 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.

More efforts have to be made in this direction by covering large representative samples, using better factorisation methods and covering different facets of achievement in terms of educational objectives. Factor structure of achievement under convergent and divergent test contents and formats is yet to be explored.

PREDICTION-ADMISSION-PROMOTION STUDIES

How effective are the predictions? There can be three approaches to handle this question, 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 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 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 the Jalota's General Mental Ability Test was a good predictor for general science ($r = 0.80$). The Nafde's Nonverbal Intelligence Test was found to be an effective predictor for mathematics. He also used other tests like the 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 he 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 a 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 intermediate level. The multiple R was 0.49.

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 subject choice at degree level by those who secured less than forty-five percent marks should be made with caution. Higher secondary examination marks in languages and mathematics were useful for selection of courses and prediction purposes. Dharm Bir and Srivastava (1965) tried to find 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, English, etc. Buch (1963) found that S.S.C.E. composite scores of physics, chemistry, and mathematics were better predictors than the individual set of scores while predicting success in the next higher examinations. Hiriyanniah (1963), while finding the prognostic value of pre-university examinations, found that 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 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 unscaled marks turned out to be as good predictors as scaled marks.

Lele and Bhagatwala (1954) developed and used the University Entrance Test for admitting students to the Faculty of Arts and Faculty of Science. The test had a maximum predictive validity coefficient of 0.74 for arts students. If the nature of students and the nature of predictor and criterion variables remain stable, prediction studies will be more useful, but unluckily, 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 should be designed and undertaken.

FAILURES

Failure in examinations is a serious problem when viewed in terms of wastage and stagnation. Education Commission (1964-66) has reported 'At present, out of every 100 children who enter Class I, only about half complete Class IV and only 34 complete Class VII'. There can be several reasons for this — economic, social and educational. In this review, there are seven studies which are directly related to the phenomenon of failures. Bokil (1956c, 1958a and 1963) conducted three studies by drawing the samples from the S.S.C.E. Board of Maharashtra. The studies by DEPSE (1964), NCERT (1964), GCPI, Allahabad (1964) and NCERT (1965) are the other four institutional studies related to the area of Failures in Examinations. Bokil (1956c) found that 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 twenty percent of the total marks. For 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 Engilsh, Bokil (1963) found that one-third of the failures were by a margin of less than ten percent marks; overall average score in English was thirty-one, which was less than the pass marks of thirtyfive. Bokil (1958a) while calculating average percentages of failures in English, found that there were significant differences from group to group and from year to year. DEPSE (1964) found that teachers' qualifications and background, teaching methods, working conditions and location of schools, transfers, buildings, equipment, clerical work done by teachers, pupils' previous attainment, pupils' attendance, media of instruction, examinations, etc., were factors which were related to pass — fail percentage of schools. It was also found that the largest failures at 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, Allahabad (1964) also found favourable results for regular candidates. It was found that girls failed less than boys. Performance in English, mathematics, science and civics was responsible for higher incidence of failure. The GCPI found that many social, economic, and educational factors were responsible for failure in schools. The NCERT (1965) conducted a sample study of failures in boards of secondary examinations taking the sample from 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.

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 characterised as inadequate, not well integrated and, therefore, at moments not goal directed.

The researches appear inadequate when the importance of the field and the existence of gaps are considered. The gaps are conspicuous by the ab-

sence 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 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 wider range of period covering 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 studies in achievement tests and evaluation procedures have been treated as an end in themselves. 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 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 tried in the tool construction. The review of the Indian studies indicates that the samples used are small and at moments not properly drawn. This creates doubts about the probability of generalisability of the findings. It is, therefore, suggested that studies in achievement testing should be designed so as to maximise the generalisability of the findings. New approaches to test construction should also be tried wherever necessary. Besides more finances and experts, this would require the use of mechanical aids like scanners (scoring machines) and computers. The

use of such mechanical devices would immediately demand the change of format, length and contents of the evaluation tools. Thus, studies for the use of mechanical aids for scoring the answersheets and declaration of examination results should be undertaken. 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 plans for the use of mechanical aids for improving the present evaluation system. Studies should also be undertaken to develop strategies for the implementation of the innovative programmes like internal assessment, grading system, question banks and national examination.

It is relevant to mention here that there is not only a scarcity of diagnostic tests in the country, but also a need to improve the diagnostic tests. The problem of failures has been a very serious one, as about half of the students fail in their final examinations. In other words, it means that half of the country's educational budget is resulting in a waste along with added frustration in the youth and society. If per chance any significant variance of this issue can be attributed to the system itself, then, we, as educators, should be held responsible for it. Such a view has already been expressed by some of the experts in the field of Educational Evaluation and Examinations. While screening the existing prediction-admission-promotion studies, four points crop up — (i) most of the prediction studies are retrospective rather than prospective; (ii) while finding success failure ratios, it is only the admitted selected sample which has 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.

The above issues, viz., achievement testing, diagnostic testing, examinations, failures and prediction, are found to be always centred around the traditional dimensions of achievement. The achievement test construction should give due emphasis to the measurement of higher mental processes as expressed by Bloom in his Taxonomy of Educational Objectives and Guilford in his SI model. The development of such tools has to be in consonance with the changed school activities, and classroom teaching and learning situation.

ABSTRACTS : 524-564

524. ASHAR, R. R., Construction and Standardization of the Diagnostic Test in Basic Algebraic Skills for the Pupils of Gujarati Medium Secondary Schools of Greater Bombay, Ph.D. Edu., Bom. U., 1972.

The study aimed at (i) standardising a diagnostic test in basic algebraic skills for the pupils of Gujarati medium secondary schools in Greater Bombay, and (ii) studying the types of errors committed by the pupils in the context of nature of units.

The study involved the usual methodology of test construction. The sample represented the population consisting of VIII, IX and X grade pupils studying in the year 1967, through Gujarati medium in Greater Bombay. Keeping in view the postal districts of Greater Bombay, schools, and grades, a sample for final study was drawn using the random clustered design. Out of the total population, 3999 boys and girls were picked up. Majority of pupils fell in the age groups of fourteen – fifteen – sixteen years. The preliminary draft of the test was administered to 268 students from five different schools. Item analysis was carried out. The final draft was having ten subtests. Different methods like test-retest, parallel form, split-half, and rational equivalence of finding reliability were employed. The median reliability coefficient was 0.90. Concurrent validity — against the external criteria of (i) the scores in mathematics at the annual examination, and (ii) the teacher's judgment — was established at all the three grades VIII, IX and X. The concurrent validity coefficients against the first criterion at VIII, IX and X grades were 0.94, 0.91 and 0.98 respectively. Norms in terms of standard scores, percentiles and stanines were established. Descriptive statistics like mean, standard deviation, percentiles, product-moment correlation, and t test were used.

The main findings of the study were : (i) norms of boys had higher values than those of girls for all the three grades VIII, IX and X; (ii) the achievement scores on all subtests of pupils in private schools were higher than those of pupils in corporation schools; (iii) sex differences in achievement scores in algebra were more prominent for pupils in private schools than for those of corporation schools; (iv) pupils' achievement was lower for two subtests of 'subtraction of algebraic terms' and 'division of algebraic terms' as compared to other subtests; (v) pupils committed errors due to lack of systematic

approach; (vi) the errors of conceptual type predominated those of the computational type; and (vii) the trend of errors continued to a greater extent in the higher grades.

525. BHOLA, V., Measurement of Achievement in Physics and Chemistry — A Critical Study of the Effectiveness of the Matriculation Examination in Physics and Chemistry conducted by the Board of School Education, Haryana, Ph.D. Edu., Kur. U., 1978.

The study had the following main objectives : (i) to study the qualitative and quantitative aspects of students' achievement in physics and chemistry of Matriculation Examination conducted by the Board of School Education, Haryana — both for morning as well as evening sessions; (ii) to analyse the different aspects of question papers contributing to the effectiveness of examination; and (iii) to compare the question papers set for the morning and evening sessions.

The answer scripts of the candidates in chemistry and physics subjects of Matriculation Examination during the years 1970 and 1971 constituted the population. Systematic sampling technique was used. Separate samples were selected for the morning and the evening sessions from the respective population of the two sessions. Four hundred answer sheets were selected from the morning session for each of the two papers and 200 from the evening session for each of the two papers in the year 1970. But in case of year 1971, a total of 950 answer sheets from the morning session for each of the two papers and 450 from the evening session for each of the two papers were selected.

The answer scripts of the candidates and question papers were analysed. The qualitative and quantitative aspects of students' achievement in physics and chemistry were analysed on the basis of their performance on individual question items, in terms of number of attempts, percentages of failures, and divisions obtained by the candidates. The general characteristics of the question items such as average score, standard deviation, frequency of the highest score and coefficient of variation were also computed. Appropriateness of question paper was studied on the basis of its format, representation of the contents of the syllabus and realisation of the objectives of science teaching. Format included choice of alternatives, possible combination of questions, allocation of marks and forms of the questions. Discriminating

The research studies should be conducted so that we have achievement tests in all the subjects with parallel forms for different regions. Suggestions of Buch (1972) of 'sequential attainment testing' should be seriously considered. A centrally organised institution aiming at conducting research on the problems of educational evaluation and examinations should be established. This institution should have extension centres in other universities and colleges. Within the present framework of the educational institutions of India, the UGC, universities, the NCERT, the CASE, State Boards of Secondary Education, and the SIE's should start providing finances and leadership to conduct studies in this area, viz., Educational Evaluation and Examinations. These are the priorities for research in this area.

power, difficulty value, reliability and validity of each question were studied. Since the study was concerned with mostly essay type question papers, some unconventional techniques were devised for the analysis of data. In some cases, techniques evolved by Gayen and others were employed.

The following were some of the major conclusions of the study : (i) Alternative question papers set for the morning and evening sessions were not analogous and of same standard. (ii) A large number of failures and very few first or second classes indicated the quantity and quality of achievement in physics and chemistry. (iii) The format of the question paper was poor from the points of content coverage and representation of the objectives of teaching physics and chemistry. (iv) The question papers were not well balanced in respect of attributes of discriminating power, difficulty value, reliability and validity of the question items.

*526. BHUSHAN, B., Experimental Verfication of Various Methods of Examination in History at the Lower and High Stages, Ph.D. Edu., Jammu U., 1978.

The major objectives of the study were : (i) to compare and contrast the efficacy of different types of examinations; (ii) to examine the reliability of scoring of examiners in different types of examinations; (iii) to indicate the relative suitability of various forms of examinations; (iv) to appraise the significant differences between internal and external examiners; (v) to find out the reliability of the grades and marks at both the stages; (vi) to look into the interrelationship between various types of examinations; and (vii) to assess overall power of discrimination of various types of examinations.

The random cluster sampling technique was employed to select the sample. Forty students from lower stage and twentyfive students from higher stage were selected. The evaluation of ten examiners each at lower and higher stages was utilised in all types of examinations. The assessment procedures employed were : (i) essay format, (ii) short answer format, (iii) open book examination, (iv) viva-voce, (v) objective type tests, and (vi) quesitonnaire regarding the preference of examiners for different types of examinations. The data were analysed by employing the statistical techniques such as measures of central tendency, coefficient of correlation, coefficient of contingency, factor analysis and analysis of variance.

The important findings of the study were as follows : (i) The internal examiner discriminated better among examiners in all these types of examinations.

(ii) There were wide divergencies in marking standards at both levels in essay type, short answer and open book type of examinations. (iii) At the higher stage short answer type tests worked well, whereas at lower stage essay type of examination worked well. (iv) Evaluation through marking or grading had equal importance at the lower stage. But the degree of correspondence was seen higher through grading at higher stage. (v) The abolition of options was favoured at both lower and higher stages.

*527. BISNAGARI, U. A., The Construction and Standardization of a Gujarati Handwriting Scale for the Pupils of Std. V to XI of Gujarat, Ph.D. Edu., Guj. U., 1976.

The aim of the present investigation was to develop a scale which could provide objective measures of handwriting quality in Gujarati language. It was intended for the use of pupils studying in standards V to XI in the schools of Gujarat State.

The content for the test was prepared by a panel of three expert teachers. The passage was administered to a sample of 2069 pupils, studying in standards V to XI of schools of Gujarat State, selected through proportionate stratified sampling. This provided 2069 handwriting specimens. Each specimen was judged by expert judges on a 3-point rating scale and ratings were converted into numerical scores. The average score given by ten judges was worked out. The specimens were arranged in the decreasing order of merit and seven representative samples were drawn covering the entire continuum and assigned the scale values 70, 60, 50, 40, 30, 20 and 10. All the specimens were scored on the basis of this scale and percentile norms were established for the different groups.

The major findings of the study were as follows : (i) The scores were distributed with a mean of 36.6 and SD 12.5. (ii) The mean scores from standards V to XI when examined, revealed an increasing trend upto standard X, but there was a slight fall for standard XI. (iii) The scores of handwriting were significantly in favour of girls. (iv) The mean scores increased upto the age of sixteen and then showed a slow decline. (v) Handwriting had a positive correlation with intelligence. (vi) The score distribution of the sample was found to be normal.

528. CHAKRABARTI, P. K., Development of a Suitable Test for Predicting Success of the School Students in Different Academic Streams, D.Phil. Psy., Cal. U., 1974.

The purpose of the study was to develop a suit-

able test in Bengali for use in predicting the probable success of students in the different academic streams : Humanities, Science, Commerce and Technical (Engineering).

The sample, for tryout, included 389 students who had passed the annual examination of class VIII. In the final standardisation, the sample consisted of 986 students including 282 boys and 231 girls from urban schools and 243 boys and 230 girls from rural schools. The test, developed in the final form, was an objective, paper-pencil and structured group test containing mostly recognition type of items. There were thirteen subtests presented in two booklets. Booklet I contained seven subtests, namely, (i) Science knowledge, (ii) Science problem, (iii) Comprehension of technical narrations, (iv) Technical problems, (v) Comprehension of commercial narrations, (vi) Computation, and (vii) Arithmetic problem. Booklet II included six subtests : (i) Synonym, (ii) Antonym, (iii) Commerce knowledge, (iv) Comprehension of mathematical principles, (v) Directed composition, and (vi) Understanding space and forms.

The coefficients of reliability found by split-half method and test-retest method ranged from 0.68 to 0.90 and 0.69 to 0.96 respectively. The validity of the test was studied subtestwise in different ways. It was validated against three sets of school examination marks obtained in (i) half yearly examination of class IX, (ii) annual examination of class IX, and (iii) half yearly examination of class X. Validity of the test was also studied by computing correlation of subtest scores with the marks obtained by 190 students studying in class XI who were adjudged as good students in their respective streams based on their examination results and their teachers' ratings. Two types of norms — percentile ranks and standard scores — were established. In order to study the nature of belongingness of the subtests, beta coefficients (coefficients of belongingness) were computed.

529. CHHAYA, M. P., Achievement in Physics of the Students of Class VIII and X of (i) The Central Schools, (ii) Public Schools of Central Board of Secondary Education, (iii) Schools of the Council of Indian School Certificate of Education, of Bombay, Delhi, Calcutta and Madras, Ph.D. Edu., Bom. U., 1978.

The major objectives of the study were : (i) to make an analytical study of the content in physics in the educational programmes for classes VIII and X of the three school systems; (ii) to construct and

standardise an achievement test in physics for classes VIII and X; and (iii) to make a comparative study of the achievement of boys and girls in physics of the schools of three systems.

The sample consisted of 1200 students of classes VIII and X selected at random from all the schools belonging to the three school systems located in the four metropolitan cities, namely, Bombay, Delhi, Calcutta and Madras. The tools prepared for the purpose of this investigation were the standardised tests of achievement. The reliability coefficients for the tests of classes VII and X were 0.903 and 0.897 respectively. The t test was employed to study the differences between different groups.

The major findings of the study were : (i) There was no significant difference between the mean achievement in physics of pupils of class VIII as well as class X belonging to Central Schools and the public schools affiliated to the Central Board of Secondary Education. (ii) The mean achievement in physics of the pupils of classes VIII as well as X of Central Schools was more than that of the pupils of schools affiliated to the Council of Indian School Certificate Examination. The difference between the two was significant at 0.01 level. (iii) The mean achievement in physics of pupils of classes VIII as well as of X of the public schools affiliated to the Central Board of Secondary Education was significantly more than that of the schools affiliated to the Council of Indian School Certificate Examination.

530. DEO, P., Effects of Revaluation on the Results of Candidates Appearing at the University Examinations, Dept. of Edu., Bom. U., 1974.

The study attempted at finding out the effects of revaluation of answer scripts on the results of the candidates appearing at the university examinations.

The sample consisted of 584 candidates who had applied for revaluation in a university in 1972. The results of revaluation of all the candidates were obtained in terms of the original and revalued marks along with the change in result. The total number of papers revalued was 815. The group of examinees was divided into AP (already passed) and F (failed) subgroups. The examinations covered were B.A. Parts I, II and III, M.A. Previous and Final, B.Sc. Parts I, II and III, M.Sc. Previous and Final, B.Com. Parts I, II and III, and M.Com. On account of revaluation the types of changes which were noted in the AP group were : (i) the merit position changed, (ii) the division changed, (iii) marks increased but otherwise no

change, and (iv) marks did not change. In the F group these categories were: (i) passed, (ii) passed with grace marks, (iii) cleared subjects, and (iv) no change in the marks and therefore no change in the original results. Mean, SD, t test and correlation were used to analyse the data.

The important findings were as under: (i) Very small percentage, viz., 2.66 of the candidates appearing for these examinations and 2.47 of the candidates appearing that year for all the university examinations applied for revaluation (AR). The largest number of AR in terms of percentages were for the candidates for final examinations, B.A. III, and B.Com. II and III. The percentage of AR in the F group was eightynine and in the AP group eleven. (ii) Out of the 584 candidates, the 'No Change' cases were 344 with a percentage of 58.9 and the cases where the results changed numbered 240 with a percentage of 41.1. (iii) The number where the result actually changed was only eighteen from the AP group and the number of those who passed due to revaluation was 210 from the F group. The 'no Change' cases in the AP group were about seventytwo percent and those in the F group were fiftyseven percent. (iv) The r between the original and revalued scores for the total group was 0.91. Considering separately for different examinations, the r's ranged from 0.80 to 0.95.

531. DESAI, S. H., Construction and Standardisation of Tests of Language Development of Gujarati Children of the Age Group 3 to 5, Ph.D. Edu., Guj. U., 1974.

The main purpose of this investigation was to study the language development patterns of Gujarati children of age group three to five. The language development patterns were to be specifically studied in terms of (a) comprehension of meaning of words and sentences, (b) sound discrimination, (c) articulation of speech sounds, and (d) oral expression of ideas.

The sample of the present study consisted of 320 school-going children selected from twentyseven schools. A test was developed to study language development of Gujarati children of the age group of three to five years.

The findings of the study were as follows: (i) The language development was highest in the beginning and slowed down in the next years. (ii) There were no sex differences in language development in the age group of three to five years. (iii) The language development of children was found to be positively related to the educational background of both father and mother. (iv) The children of rich parents showed better language development than those of poor parents. (v) The size of the family was also found to have a positive effect on the language development of children.

532. GAUR, R. N., Standardisation of Achievement Tests in Different Aspects of Hindi for Matriculation Students of Haryana, Ph.D. Edu., Kur. U., 1973.

The main objective of the study was to construct and standardise an achievement test in different aspects of Hindi for matriculation students of Haryana.

The six tests included in the battery covered the following aspects of Hindi: (i) spelling, (ii) vocabulary, (iii) applied grammar, (iv) formal grammar, (v) idioms and proverbs, and (vi) comprehension. About one and a half times more items than proposed to be kept in the final tests were constructed after consulting textbooks, syllabi, examination papers, exercise books and experienced teachers. The items were tried out on a sample of 370 students of matric class of Haryana. The final draft was administered to 1,500 students of matric class belonging to seven districts of Haryana selected by cluster sampling technique. Percentile rank, percentile band, and T-score norms were established for each test.

The reliability of each test was established by split-half method. The reliability coefficients for the tests varied from 0.83 to 0.92. The validity of the tests was assumed as they were achievement tests and had been constructed after consulting the syllabi, prescribed books, supplementary readers, the examination papers, teachers, and sample exercise books of local students. No criterion was available for external validation since marking was not done separately on different aspects in examinations and teachers were not having a clear idea of the proficiency of their students aspectwise.

533. GAYEN, A. K., NANDA, P. B. and others, Measurement of Achievement in Mathematics, IIT, Kharagpur, 1961. (MOE financed)

This survey was undertaken with an objective to study statistically the effectiveness of Board examinations with reference to mathematics, and to suggest in what respect the then existing system of examination had to undergo change. In this context the study aimed at (i) analysing the nature of paper setting of school final mathematics

paper; and (ii) making suggestions for improving paper setting.

Considering each examination centre as a stratum, a stratified random sample of 3,717 answer scripts in compulsory mathematics of the first school final examination of the year 1952 of the West Bengal Board was selected. The nature of paper setting was analysed on the basis of a statistical study of the performance of students in the paper. Comparison of marks and percentages of failures, passes and non-attempts with various question items and their alternatives was made. An investigation into the reliability and validity of the question paper was done.

Some of the findings of the study were: (i) the examination in mathematics did not so much measure mathematical ability; (ii) some of the important topics were not covered in the question paper; (iii) questions which really measured mathematical ability were poorly answered; (iv) the questions lacked discriminatory value, with most of the questions discriminating only at low levels; and (v) questions of unequal difficulty value were set as alternatives.

*534. GIRI, G. P., Construction of Vocabulary Test in Hindi for the Purpose of Classes VI Through IX, Ph.D. Psy., Pat. U., 1976.

The objective was to construct a test of Hindi vocabulary to assess the vocabulary development of the children.

Five pupils from each of classes VI through IX of twentynine schools were asked to write down as many Hindi words as they could. This made a pool of 1200 words which was administered to a sample of 400 pupils selected from the specified grades of the same twentynine schools. The t test and analysis of variance were employed for data analysis. Two hundred items were retained and these were administered to 800 pupils with 100 boys and 100 girls from each class, selected at random. Item analysis was done and on that basis 100 items were retained. This final test was administered to 300 pupils taking seventyfive pupils from each class.

The coefficient of reliability calculated by Kuder Richardson formula was 0.98. The concurrent validity was calculated by correlating the scores of each class with examination marks and the Mehta's Intelligence test. The coefficients of correlation between the test score and the intelligence test ranged from 0.66 to 0.89 and those between the present test and examination marks varied from 0.55 to 0.81. The age norms, grade norms and percentile norms were established for this test.

*535. GNANAPRAGASAM, N. S., Setting-up of a Question Bank of Objective Tests, Yeddanapalli Institute of Educational Research, Loyola College, Madras, 1975. (UGC financed)

The present project was undertaken as the College Science Improvement Programme (COSIP) initiated by the UGC. The main purpose of the project was to set up a question bank as a practical and significant step towards examination reform. Necessity to switch over to objective tests as a better measure of evaluation was felt on the following grounds: (i) Since a large number of questions can be included in an objective test, the syllabus can be covered adequately. (ii) Marking is rapid, reproducible and reliable in such a test.

For developing questions or items for the bank, objectives of testing were taken to be knowledge, comprehension, application, analysis, synthesis and critical evaluation. Of all the different types of objective test items, multiple choice items with three to five choices were largely developed in this project. Eleven other types of items included the following: multiple completion, master list, statement and reason, true or false statement, matching, filling up blanks, classification, comparison, simple problems, short answer questions, and comprehension passage. Provision was made in the marking scheme for minimising guessing error. The draft questions were discussed with the staff handling the particular subject and necessary screening was done. Each question was then typed on a card and arranged subjectwise in a card filing cabinet. For the sake of tryout, test papers were prepared by selecting questions out of this pool of items and then administered to the preuniversity and degree students of science in Loyola College, who formed the sample of the study. Item analysis was carried out. Facility index and discrimination index were found out for each item which would help the teacher to pinpoint the suitable questions in the objective test at the time of selecting them for preparing test papers, in terms of difficulty level, consistency and relevance of the questions and effectivity of the distractors.

The most important outcome of the project is the creation of Loyola Chemistry Question Bank. The bank consists of objective type of questions, mostly at the B.Sc. level, numbering 2983, of which 1847 have already been tested during thirtyeight objective tests that were conducted throughout the project period. The questions have been classified into various topics in the three main branches of

chemistry, namely, organic, inorganic and physical.

536. *GOHIL, H. B., Construction and Standardization of a Silent Reading Test in Gujarati for Pupils Studying in Standard VII in the Schools of Saurashtra, Ph.D. Edu., Sau. U., 1974.*

The present study was conducted to fulfil the following objectives : (i) to construct a Silent Reading Test in Gujarati so as to cover three different aspects of silent reading skill that were in tune with the pupils' age group; (ii) to assign three different subscores, viz., those on reading rate, reading comprehension and vocabulary, in addition to the total score that was given to the pupils; (iii) to prepare Z-scores, H-scores, T-scores, C-scores and stanines in order to make the pupils' raw scores meaningful and interpretative; and (iv) to design a pupils' self interpreting reading profile so as to assist the primary school leavers by offering developmental guidance.

Silent Reading Tests were designed with consideration for : (i) reading rate; (ii) reading comprehension; (iii) sentence meaning including word meaning, usage of proverbs and idioms; and (iv) location of information. Eight subtests were constructed for the battery of tests. They were on : (i) reading rate; (ii) story comprehension; (iii) prose comprehension; (iv) poetry comprehension; (v) paragraph comprehension; (vi) word meaning; (vii) vocabulary; and (viii) idioms and proverbs. Test was given to 370 students for tryout. Test items were revised, redefined and selected on the basis of item analysis. In the final form of the test, there were ninetyeight items divided into eight subtests. In all, 6,000 pupils were selected from 194 schools of Saurashtra for the final administration of the test.

Norms for grade VII for boys group, girls group and the whole group were established. Measures of skewness on the basis of median and percentile were 0.099 and 0.039 respectively and the measure of kurtosis was 0.274; the normality of the distribution was also tested by using the method of moments. The reliability coefficient was found out by using six different methods and techniques, and it ranged from 0.89 to 0.96. The validity coefficient of the tests fell between 0.56 and 0.77. The general findings of the factor analysis corroborated the aspects on which the test had been constructed. Factors concerning age, caste, parents' literacy, parents' income and the area where the schools were situated, were found to affect the silent reading ability of the pupils.

537. *GUJARAT RESEARCH SOCIETY, Construction and Standardisation of Achievement Tests in Gujarati for Standards V, VI, and VII, Bombay, 1963. (MOE and NCERT financed)*

The aim of this project was to construct and standardise achievement tests in Gujarati, Hindi, arithmetic, history, geography and science for classes V, VI, and VII.

The following three types of tests were constructed : (i) recall—simple recall and completion, (ii) recognition—multiple choice, matching and true false type, and (iii) figure type. A sample of 100 students was taken for the purpose of pretryout of the test. For the pilot testing, the sample consisted of 370 students for each test. For the final study, the sample comprised 939 boys and 640 girls from thirtyfive Gujarati medium schools of Bombay city. Correlation was used as the statistical measure for analysing the data. Age norms, grade norms, percentile norms and standard score norms were also worked out.

It was found that (i) the split-half reliability coefficient of the tests ranged from 0.79 to 0.99, and (ii) the validity coefficient of the tests ranged between 0.21 and 0.77.

538. *GUPTA, S. K., Factor Analysis of Attainments of Higher Secondary/Pre-University Passed Students in Different Aspects of Physical Sciences and Mathematics, Ph.D. Edu., Kur. U., 1974.*

The study aimed at investigating the factorial structure of attainments in physical sciences and mathematics for pupils of higher secondary|pre-university stage.

In the study the descriptive correlational methodology was employed. The sample consisted of 200 students belonging to the state of Haryana. These students had passed the higher secondary or pre-university examination. The sample was drawn from colleges situated at district headquarters and also interior places of the districts. A battery of the following nine tests was developed : (i) Physics Test 1 (facts and technical terms), (ii) Physics Test 2 (principles and their applications), (iii) Physics Test 3 (numericals), (iv) Chemistry Test 1 (facts, symbols and formulae, and equations), (v) Chemistry Test 2 (chemical laws, properties and their applications), (vi) Chemistry Test 3 (numericals), (vii) Mathematics Test 1 (algebra), (viii) Mathematics Test 2 (trigonometry), and (ix) Mathematics Test 3 (geometry). All these tests were having multiple choice items. The number of items in all the tests ranged from thirty to thirtyeight. The split-half reliability coefficients for these tests ranged

from 0.78 to 0.95 with a median value of 0.83. The tests were supposed to have content validity. Data were collected and converted to T scores whenever Kolmogorov-Smirnov Test indicated 'not normal' distribution. Data were analysed by employing product-moment correlations, Thurstone's centriod method of factor analysis, and oblique rotation with the method of extended vectors.

On perusal of the psychologically meaningful factor matrix and the matrix of correlations between the oblique factors, the following factors were discovered : (i) General Mathematical Ability Factor having high loadings on all the three mathematics tests; (ii) General Factors of Scientific Ability having high loadings on all the tests of physics and chemistry; (iii) Symbol and Number Factor having high loadings on Physics Test 3 (numericals), Chemistry Test 3 (numericals) and Mathematics Test 1 (algebra); (iv) Deductive Reasoning Factor having high loadings on Physics Test 2 (principles and their applications) and Chemistry Test 2 (chemical laws, properties and their applications); and (v) Space Factor having significant loadings on Physics Test 3 (numericals) and Mathematics Test 2 (trigonometry).

539. HIRA DEVI SUJAN SINGH, Construction and Standardisation of Achievement Tests in General Science for Standards V, VI and VII for Children studying through Sindhi as the Medium of Instruction in Greater Bombay, Ph.D. Edu., Bom. U., 1973.

The objective of this study was to construct and standardise achievement tests in general science for standards V, VI and VII for children studying through Sindhi medium in greater Bombay.

The pre-tryout test was administered to a sample of 100 students from each of the standards V, VI, and VII. The sample was drawn from schools situated in different localities and included students from all strata of the society. The tryout form of the test was administered to a random sample of 185 students in each of the standards V, VI and VII. The selection of the items for the final run was based on the specifications of the contents and the difficulty value and discriminating power of each item. The final form of the test was administered to 410 pupils of standard V, 380 pupils of standard VI, and 440 pupils of standard VII.

The reliability coefficients computed by different methods for the three tests were as follows : (i) test-retest method — ranging from 0.92 to 0.96; (ii) split-half method — ranging from 0.96 to 0.97; (iii) K–R formula — ranging from 0.86 to 0.90;

and (iv) Rulon's formula — ranging from 0.88 to 0.91. The validity of the tests was found to be (i) between 0.94 and 0.98 by correlating test scores with the annual examination marks, and (ii) between 0.88 and 0.93 by correlating test scores with the teachers' rating (rank correlation method). Stanines, percentiles, standard scores and T-scores were worked out.

540. I. T. COLLEGE, Construction and Standardisation of a Diagnostic Reading Test in Hindi for Class III Primary, Lucknow, 1971. (NCERT financed)

The project aimed at constructing and standardising a diagnostic reading test in Hindi for class III.

Items were constructed and were run as a pilot study on a sample of fiftytwo pupils of class III. One hundred items were selected and divided into five parts. This form of the test was administered to a sample of 327 pupils of class III of schools of six selected areas of U.P. The final form of the test consisted of five parts, viz., Hindi usage – giving alternative, Hindi usage – filling the missing words, vocabulary, paired combinations, and silent reading. The final test was administered to 937 pupils (572 boys and 365 girls), both rural and urban, of class III of schools of U.P.

Concurrent validity of the test was found to be 0.69. The coefficient of reliability by split-half method was found to be 0.90. Percentile norms were developed.

541. JHA, J., Development of a Battery of Tests for measuring some of the Basic Skills of Students of Classes VI and VII, Ph.D. Edu., Pat. U., 1974.

The aim of the study was to construct and standardise a battery of tests to measure and evaluate functional mastery of some of the basic skills of the pupils of classes VI and VII. The four dimensions were : (i) arithmetic concepts, (ii) concepts in problem solving in arithmetic, (iii) understanding of basic concepts, vocabulary, definition and facts of geometry, and (iv) written language skills involving formal factors such as spelling, grammatical factors such as common errors in language form and sentence structure and form, and such elements of composition which could be relatively tangible, objective and measurable.

The difficulty values as well as discrimination values of the sets of items, after the tryout, were taken into consideration for selecting the test items for the final form. Finally, four pools of 30, 18, 12

and 50 items were selected for class VI, and 22, 12, 16 and 50 for class VII. The test-retest reliability coefficients ranged between 0.784 and 0.933 for class VI and between 0.760 and 0.864 for class VII. The split-half reliabilities of the full length test after the application of the Spearman – Brown formula for class VI ranged between 0.834 and 0.929 and for class VII it ranged between 0.363 and 0.965. The coefficients computed by the K-R formula ranged between 0.820 and 0.940 for class VI and between 0.812 and 0.930 for class VII. The tests were validated against the Mohsin's Intelligence Test — II and school examination marks (an average of two consecutive examination marks). A unified set of percentile norms each for class VI and class VII for boys and girls of rural and urban settings were established.

542. *JHAVERI, B. J. and PATEL, B. C., A Study of Inter and Intra Examiner Reliability in Marking Essays with and without using the marking scheme, New Arts College, SPU, 1968. (NCERT financed)*

The study aimed at testing two hypotheses, namely, (i) there would be significantly higher intra-examiner reliability in the case of marking an essay with a scheme than marking without a scheme; and (ii) the intra-examiner reliability would be significantly higher in the case of well defined essays than the traditional essays.

With the help of a workshop organised under the joint auspices of the NCERT and the SSCE Board of Gujarat State, different types of essays were prepared, and listing of different abilities required for effective essay writing at the pre-university level were decided. Two test forms were administered and collected from eightythree students of preparatory arts classes. Four examiners checked twice the same answerbooks under two conditions — without marking scheme and with marking scheme. This procedure of marking was followed for well defined essays and the traditional essays.

It was found that (i) the use of marking scheme significantly increased the mean assessment; (ii) the use of marking scheme did not increase any type of reliability; and (iii) the reliability of the assessment of well defined essays was found to be greater than that of traditional essays.

543. *MODI, D. J., Construction and Standardisation of Listening Comprehension Test in Gujarati for Class VIII, Ph.D. Edu., MSU, 1975.*

The main objectives of the study were : (i) to standardise a test for assessing listening comprehension; (ii) to find out the common variance between the factors of listening comprehension and listening ability; (iii) to find out the sexwise differences in achievement on the constructed test; and (iv) to study the rural-urban and culture (area or district) influence on the achievement of pupils.

Five factors, namely, vocabulary, attention, memory (meaningful and rote), auditory resistance, and reasoning (inductive and deductive), were selected. For the preliminary tryout a battery of thirtyeight subtests was constructed involving these factors. Two different listening situations, namely, oral presentation and tape-recorded presentation, the latter having an equal percentage of female and male voices, were developed. The preliminary tryout was made on 100 students. Item analysis was carried out. The pilot test battery which had 701 items consisted of thirtyeight subtests grouped under five main tests, namely, (i) listening to different situations (fourteen subtests) and one vocabulary test, (ii) attention tests (five subtests), (iii) memory tests (seven subtests), (iv) auditory resistance tests (five subtests), and (v) reasoning tests (six subtests). The battery was administered to a sample of 466 pupils including both boys and girls from urban and rural areas of the standard VII in thirteen high schools of five districts. The items were subjected to item analysis. The finally selected subtests were divided into two parallel forms A and B, each consisting of 180 items. The final forms of the test were administered to a sample covering boys and girls from rural and urban areas and the five main areas of culture in Gujarat. The sample consisted of 1905 pupils for Form A and 1943 pupils for Form B.

The coefficient of reliability by test-retest method was found to be 0.80 and by parallel form method it was found to be 0.80. The correlation coefficients of the subtests with the total test ranged from 0.27 to 0.65, which were significant at 0.01 level. The factor analysis rendered vocabulary, memory, auditory resistance, attention and reasoning as the components of listening comprehension. The developed test correlated 0.63 (Form A) and 0.71 (Form B) with the Desai-Bhatt Group Test of Intelligence and 0.35 (Form A) and 0.40 (Form B) with the annual examination scores in Gujarati of standard VII. It was also found that the socio-economic status of the pupils influenced the listening comprehension, the results being in favour of higher status (significant at 0.05 level). There was no significant difference between boys and girls in the performance on the test, whereas a significant difference was found

between the performance of rural and urban students.

544. Md. FAIZUL ISLAM, Construction and Standardisation of the Achievement Test in General Science for Students of Class VII of Bihar, Ph.D. Edu., Pat. U., 1975.

The purpose of the study was to construct and standardise an achievement test in general science for students of class VII in the State of Bihar.

After specifying the objectives of teaching science in schools, content analysis of the course was done. The preliminary draft of the objective type test was administered to 264 students of class VIII in two different schools of Ranchi. The students were just promoted from class VII to VIII. The test consisted of 104 items, but after the item analysis 18 items were rejected. The final form of test, having 86 items was standardised on a sample of 300 boys and 300 girls reading in class VIII of selected schools of Bihar State. The scores were statistically analysed for finding out skewness and kurtosis. Standard score norms, T score norms and percentile norms were computed for male and female samples separately.

545. NATH, B., Inter-Zonal Analysis of P. U. Results, 1969, Examination Research Unit, Gau. U., 1972. (UGC financed)

The present study aimed at finding out the effect of scaling in different subjects.

The area under the jurisdiction of the Gauhati University was divided into three administrative zones, namely, Karimganj, Tezpur and Gauhati. Four major elective subjects, namely, economics, history, political science and education were taken into account. A total of 33,778 answer scripts of the P.U. (A) examination of 1969 were randomly selected, the randomisation being followed zonewise. Using medians as bases, the marks were scaled.

It was found that (i) the effect of scaling was more or less uniform over all the subjects in the different zones; (ii) scaling reduced the dispersion of mark distributions thus bringing down the variations between the examiners; and (iii) there were some real differences between the different zones in terms of student performance.

546. NATH, B., Correlation between the Marks of Two Successive Examinations, Examination Research Unit, Gau. U., 1973. (UGC financed)

The investigation was undertaken to find the extent to which examination results were reliable indicators of performance in successive examinations. The examinations selected for the study were the Previous Examination, 1966 and the Final Examination, 1967 of the Master Degree courses of the Gauhati University in the subjects of English, economics, and political science. The students who passed the Previous Examination in 1966 in the three subjects under consideration formed the sample and this happened to be thirtynine, seventytwo and ninetyone respectively. The percentage of marks of these students in both the Previous and the Final Examinations were noted. Regression equations for all the three subjects were developed. Correlation coefficients between the marks of Previous and the Final Examinations were found out for English, economics, and political science separately.

The coefficients of correlation for English, economics, and political science were found to be 0.76, 0.82, and 0.83 respectively.

547. NATH, B., A Comparative Study of the Results of Some Selected Colleges in the P. U. (A) Examination, 1971, of the Gauhati University, Examination Research Unit, Gau. U., 1974. (UGC financed)

The main objective of the present study was to ascertain the factors influencing the examination results and to suggest certain measures for their improvement.

Out of the total of about seven thousand candidates who appeared in the Preuniversity (Group A) Examination in 1971 of the Gauhati University, a representative sample of 601 candidates was drawn. It happened that out of eightynine colleges, only fortythree were represented in the sample. The marks secured in the major subjects — English, education, economics, geography, history, and political science — by the candidates in the sample were tabulated collegewise. For each subject the data were analysed by applying F test. Further by applying the principle of gradation charts, the colleges were classified for each subject into three categories, namely, below average, average, and above average in standard. The quality of the students admitted into the preuniversity of classes was noted. Plotting of student—teacher ratio of a college in a subject against the percentage of pass gave a bivariate distribution. The correlation coefficients and the regression equations for all the subjects under study were calculated.

It was found that (i) the highest number of colleges in the below average category was in the subjects of economics, English and geography; (ii) the poor standards in results were mainly due to two factors — (a) poor quality of students admitted, and (b) high student—teacher ratio.

*548. PALSANE, M. N., Predictive Potentiality of the Secondary School Certificate Examination Scores, M. B. Patel College of Education, SPU, 1965. (NCERT financed)

The aims of the study were (i) to examine how far the total performance of the candidates at the S.S.C. Examination correlates with the total performance at the subsequent university examination; and (ii) to see how far the performance in specific or groups of subjects at the S.S.C. Examination correlates with the total performance at the subsequent university examination, as well as performance in specific subject or groups of subjects.

The sample of the study involved two batches of students. The first batch consisted of students who had appeared at the S.S.C. Examination in 1957 and had entered the university the same year in any of the faculties of agriculture, arts, commerce, and science. The second batch consisted of the students who appeared at the degree examinations in 1963 in the faculties of agriculture, arts, commerce, science, and engineering. The data related to age, sex, caste, rural-urban origin, father's occupation, number of subjects offered and the marks obtained at the S.S.C. Examination were collected. Those students who passed and appeared in the examination in the subsequent years were followed up. The students of the second batch were traced backward upto their pre-university examination in their corresponding faculties. Product-moment correlation was used to analyse the data.

The study revealed the following : (i) the predictive value of the grand total marks at the S.S.C. Examination was found to be quite high; (ii) languages group marks had highest predictive potentiality in the faculties of arts and engineering; (iii) in the faculty of commerce, marks in social studies at the S.S.C. Examination appeared to be a good predictor variable; (iv) in the faculties of engineering and science, marks in mathematics at the S.S.C. Examination were found to be a significant predictor variable; (v) marks in science subjects at the S.S.C. Examination had high predictive value in the faculty of engineering, but poor in the faculty of science; and (vi) in the faculty of arts, performance in the social studies at the S.S.C. Examination was a moderate predictor.

549. PANDYA, K. D., Standardisation of a Language Ability Test in Gujarati for College Entrants, Ph.D. Edu., MSU, 1973.

The purpose of the study was to standardise a test of language ability for students entering a college.

The test consisted of four parts, viz., (i) vocabulary (eight areas); (ii) sentence structure (three areas); (ii) spelling and punctuation (three areas); and (iv) comprehension (three areas). The initial tryout test contained 584 items. After item analysis, 400 items were retained. Standardisation of the test was done on a sample of 620 boys and 380 girls studying in preuniversity classes of Gujarat, South Gujarat and Saurashtra universites.

The reliability coefficient of the test by using test-retest method was found to be 0.82 (N = 100). The study yielded a correlation coefficient of 0.69 with the Uravasi Desai's Language Ability Test for high school students. The test also had a correlation coefficient of 0.75 with the marks obtained by the students in Gujarati language subject in their S.S.C. Examination. Intercorrelations between the subtests ranged from 0.43 to 0.58.

*550. PATEL, J. M., Preparing State Norms for Delta Class (Standard VIII) Language (Hindi), History and Geography, Ph.D. Edu., Bom. U., 1977.

The objectives of the study were : (i) to prepare achievement tests in Hindi, History and Geography for the delta class; (ii) to standardise the achievement tests in Hindi, History and Geography prepared for the delta class; (iii) to determine state norms for pupils on the basis of newly prepared tests in Hindi, History and Geography for the delta class; (iv) to study intergroup achievement in the subjects; (v) to judge the quality of the developed tests; (vi) to provide schools with a reliable and valid tool for measuring their pupils' achievement in Hindi, History and Geography for the delta class.

Three tests were prepared on the basis of the syllabus and textbooks published by the Department of Education, Government of Gujarat for the classes V, VI and VII.

The findings of the study were : (i) pupils were able to achieve only fifty percent or even less than that; (ii) memory, practice and test familiarity did not affect the test scores of all the three tests; (iii) the mean difference between urban boys and urban girls was not significant; (iv) pupils of urban areas were high achievers; (v) rural girls were superior to rural boys in all the three tests; and (vi) there was a direct relationship between achievement and age.

*551. PATEL, T. M., *Development of a Battery of Diagnostic Tests in Arithmetic for the students studying in Standards V, VI, and VII in Gujarati Medium Schools in Greater Bombay, Ph.D. Edu., Bom. U., 1976.*

The objectives of the study were : (i) to identify the pupils who were deficient in arithmetic; (ii) to discover the areas of difficulty encountered by the pupils; (iii) to give remedial teaching based on the analysis of errors committed; and (iv) to measure the outcomes of remedial teaching.

Test items were constructed and different subtests were prepared. A pilot study was conducted and items were rearranged on the basis of difficulty value. The final tests were given to 6190 pupils from nineteen schools of Greater Bombay.

Two criteria were applied to select the pupils for pre-remedial tests, viz., low scores on diagnostic test and failure in mathematics in the annual examination. Remedial teaching was followed by classroom teaching. Post–remedial teaching tests were administered as soon as the remedial teaching was over, and the gain caused by the remedial teaching was computed. The study has offered some suggestions and recommendations for improving the standard of achievement in arithmetic.

*552. SALI, V. Z., *Construction and Standardisation of Unit Tests in Physis for Pupils of Standard VIII, Ph.D. Edu., Poona U., 1977.*

The objective of the study was to construct and standardise five unit tests in physics for standard VIII for Maharashtra State Board Secondary Schools.

The sample for each unit test during the pilot study ranged from 107 to 149 students (both boys and girls) in standard VIII drawn from sixtysix schools in Poona district. Item analysis was carried out. The final form of the study was administered on a sample constituting 6130 students studying in standard VIII of thirtyone schools in Poona district.

Norms were developed in the form of percentiles, stanines, Z-scores and T-scores. Reliabilities obtained through K-R 21 ranged between 0.65 and 0.79 for the five unit tests. By Stanley's formula, the r's ranged between 0.68 and 0.83. Content validity was determined on the basis of internal consistency of the test. Validity coefficients ranging between 0.69 and 0.80 were obtained when validated against teachers' assessment in terms of ranks.

553. SHARMA, N. R., *Examination Reform — An Analysis of Public Examination Results of U.P. Board, Government Higher Secondary School, Aliganj, New Delhi, 1966. (NCERT financed)*

The major objectives of the study were : (i) to find out the main characteristics of the distribution of scores obtained by regular students at the high school examination in English, mathematics, Hindi, and science; and (ii) to find the effect of change in syllabus on the pupil performance.

The sampling involved the stratified random sampling technique, the stratification being done with respect to two variables, namely, sex and region (rural and urban). Thus a sample of 5,000 scores was taken for different years in three compulsory subjects (Hindi, English, and mathematics). Mean, median, and SD were calculated. Chi-square test was applied and significance of difference between means was also found out.

The study revealed that (i) from year to year the average performance continuously increased in Hindi, mathematics and science, but in English a reverse trend was noticeable from 1961 to 1965; (ii) the reduction on curricular load in 1964 led to significantly better average performance in all the individual subjects taken for study; (iii) the rural candidate on an average got five or more marks less in every subject than his urban counterpart; and (iv) the correlation coefficients between home and public examinations were found to be positive and significant at 0.05 level.

554. SHARMA, S. P., *Biological Science Aptitude Test for School Students, Ph.D. Edu., Pat. U., 1975.*

The study aimed at developing a test for measuring aptitude for biological sciences of school students of higher secondary stage, taking into account the conditions prevailing in the country.

In order to identify the areas relevant enough to be included in the biological aptitude test, twenty-five judges consisting of higher secondary biology teachers, university professors of bio-sciences, doctors, agricultural scientists, psychologists and educationists, were asked to select areas on the basis of their importance and relevance from a list of twenty areas. The judges were also asked to add as many additional areas as they thought fit for inclusion in the test. On the basis of hundred percent agreement, the selected six areas were verbal reasoning, numerical reasoning, biological science comprehension,

sympathy towards living being, experimental bent, and observation and interpretation. Covering these areas a list of 224 items was prepared. After screening by a group of ten biology teachers 131 items were retained and included in a preliminary form of the test which was administered to 100 students of class IX in two higher secondary schools of Delhi. Item analysis was done by using biserial correlation. Rejecting the items with low values of biserial r, the final form of the test comprised seventy items. The test was then administered to a sample of 500 class IX students of higher secondary schools of Delhi. The distribution of the final scores, when tested for normality, did not depart significantly from normality. Percentile and T-score norms for the test were worked out.

It was found that the reliability coefficients of the test by the mehods of test-retest, split-half and K—R formula 20, were 0.96, 0.80 and 0.82 respectively. Predictive validity of the test was estimated by correlating the scores of 100 class IX students on this test with their marks obtained in biology at the annual examination. The coefficient of validity was found to be 0.77 which was highly significant.

*555. SHARMA, V. P., Achievement Norm Study of School Children of Ahmedabad, B. M. Institute of Mental Health, Ahmedabad, 1977. (NCERT financed)

In this study, an attempt has been made to explore achievement of children in relation to population variables, namely, (i) the school system and (ii) the sex of the children.

Sample included corporation schools and recognised schools. In all, 1027 children, 503 from corporation schools and 524 from recognised schools, were selected. Shah's rating scales for schools (for A and B grade schools) were administered in a limited geographical area of Ahmedabad.

The following were the findings of the study : (i) No significant differences were observed in the achievement of boys and girls in arithmetic and in copying of designs, but in reading test the girls achieved significantly higher scores, the significance being at 0.02 level. (ii) When the scores of corporation and recognised schools' children were compared, highly significant differences were observed. The children of recognised schools were found to have achieved significantly higher scores in reading, arithmetic and copying of designs, as compared to those of corporation schools. These differences for all the three tests were significant at 0.001 level.

*556. SHARMA, V. S., Battery of Tests for the Delta Class in General Science and Mathematics (Analysis, Validation and Standardisation), Ph.D. Edu., Udai. U., 1976.

The chief objective of the present study was to construct and standardise a battery of tests in mathematics and general science for the delta class pupils studying in different parts of Rajasthan. A few subsidiary objectives were as follows : (i) To find out the specific objectives in terms of concepts, understanding, application and skills which were envisaged by the teachers and the curriculum planners for class VIII in mathematics and general science; (ii) to make a comparative study of the syllabi in mathematics and general science in Hindi speaking states; and (iii) to compare the standards of achievement in both the subjects in respect of sex and rural-urban strata of the population and ditsricts of the state.

For the pre-tryout of the test, seven pupils (two brilliant, three mediocre and two dull) were selected. For the tryout of the test, 500 pupils from eleven institutions were taken and for the final form of the test 1708 pupils were selected from fortyeight institutions. The reliability of the tests was computed on a sample of 400 pupils (200 boys and 200 girls) by Guttman and Kuder-Richardson formula using split-half method. For the validity of the tests, (i) internal consistency, (ii) content validity, (iii) concurrent validity and (iv) convergent validity were ascertained.

The findings of the study were as follows: (i) It was found that the reliability coefficients of the tests for general science and mathematics were ranging from 0.918 to 0.996. The indices of reliability in case of general science and mathematics were found to be 0.959 and 0.989 respectively. The validity coefficients were ranging from 0.41 to 0.53. The girls population in the state scored highest in general science as compared to the other strata of the sample, when means were calculated; the urban strata of the sample scored second highest and the boys strata the lowest in comparison with the state level mean score. In the case of mathematics too, the girls scored highest, the urban population occupied the second position while the boys strata retained the lowest position. (ii) The performance of the pupils in general science was highest in the districts—Sirohi, Sikar and Tonk and the lowest in Bikaner, Udaipur and Bundi districts. The performance of the pupils in mathematics was the highest in the districts of Alwar, Ajmer and Sirohi while it was found to be the lowest in Bundi, Sawai Madhopur and Udaipur districts. (iii) The girls were found to be significantly superior

to boys both in general science and mathematics. There was a significant difference (at 0.05 level) between the performance of the rural and the urban population in general science, whereas there was no significant difference between the performance of the rural and the urban population in mathematics.

557. SIE (Kerala), Construction and Standardisation of an Achievement Test in General Mathematics for Standard VIII Students in Kerala, 1965. (NCERT financed)

The study aimed at standardising an achievement test in general mathematics for pupils of standard VIII in Kerala.

The preliminary form of the test included ninety items in the areas, namely, (i) skills, (ii) facts, terms, and concepts, and (iii) application. For the pilot study, there were eightyfive items which were administered to a sample of 500 pupils belonging to twelve schools of Trivandrum and Attingal districts. After item analysis, sixty items were included in the final test. For the standardisation, the test was administered to a sample of 1,400 pupils (760 boys and 640 girls) of fortyeight schools belonging to different districts both of rural and urban areas.

The reliability coefficient calculated by split-half method and corrected by Spearman-Brown formula was 0.88. The empirical validity coefficient of the test was found to be 0.67. Standard scores, percentile scores, and normalised T-scores were computed.

558. SIE (Kerala), Construction and Standardisation of an Achievement Test in General Science for Standard VIII Students in Kerala, 1965. (NCERT financed)

The study aimed at standardising an achievement test in general science for grade VIII students of Kerala.

Based upon the analysis of content courses and general objectives of science teaching, the test items related to (i) recognition of facts and information, (ii) skill in problem solving, (iii) ability to discover, (iv) ability to reorganise, and (v) skill in manipulation of instruments, were constructed. The test consisted of 156 items—ten completion type, fiftysix multiple choice type, four true and false type, twentysix matching type, thirtynine classification type, and nineteen pictorial type. The pilot test sample consisted of 500 pupils of grade IX at the beginning of the year. The standardisation sample consisted of 915 boys and 485 girls of urban and rural areas of Kerala.

The test had a split-half reliability coefficient of 0.67. The concurrent validity coefficient against the criterion of the teacher-made test was found to be 0.83. Standard scores, normalised T-scores and percentile norms were also computed.

559. SIE (Kerala), Construction and Standardisation of an Achievement Test in Social Studies for Standard VIII Students in Kerala, 1965. (NCERT financed)

The study aimed at standardising an achievement test in social studies for standard VIII.

The items were constructed keeping in view the knowledge and application objectives. The items on knowledge objective covered (i) recollection of facts and principles, (ii) recognition of illustration of facts, (iii) discrimination, (iv) classification of facts and phenomena, (v) comparing and contrasting, and (vi) giving symbolic representation. The items on application objective covered (i) analyses, (ii) inference, (iii) explanation, (iv) establishing relationship, (v) judgment of adequacy or inadequacy, and (vi) location of information. The pilot run of the test was conducted on a random sample of 500 pupils studying in standard IX of schools in Kerala state. The final form of the test consisted of 150 items of which thirtyfive were of completion type, fifty of matching type, sixteen of classification type, thirtyseven of multiple choice type and twelve items were based on maps. The final form of the test was administered to a sample of 1,400 pupils (911 boys and 489 girls).

The coefficient of reliability by the split-half method as corrected by the Spearman-Brown formula was found to be 0.72. The test was validated against a teacher-made test and was found to have a coefficient of correlation of 0.96. Standard scores, normalised T-scores and percentile norms were developed.

*560. SINGH, A., Construction of a Battery of Objective Tests for Assessment of Proficiency in Writing English Composition, Ph.D. Edu., Kur. U., 1978.

The main objective of the investigation was to assemble a battery of objective tests for the assessment of proficiency in writing English composition, of high school finishers in Hindi speaking regions.

The battery of tests was constructed on different aspects of written English composition, namely, spelling, punctuation, vocabulary (phrases), vocabulary (words), paragraph organisation, applied grammar, general knowledge, and handwriting. Except the handwriting scale, item validity was calculated for all the tests through biserial correlation on the basis of Flanagan's tables. The items with difficulty values rang-

ing from 0.30 to 0.70 were selected for the tests. The final draft of the tests was administered to 500 students for establishing split-half reliability. The amount of time taken by about ninety percent of the students to complete the tests was fixed as the duration of time for different tests. The handwriting scale was evolved by Thurstone's Equal Appearing Interval method. For the final study 245 rural students and 255 urban students of pre-university classes (PUC) were drawn from the colleges affiliated to Himachal Pradesh University. For the purpose of identifying the crucial variables for predicting college students' proficiency in writing English composition, the step-wise regression was carried out.

As the variable of general knowledge increased the multiple R by .00017 only, it was discarded. Finally, the battery consisted of seven tests on spelling, punctuation, vocabulary (phrases), vocabulary (words), paragraph organisation, applied grammar, and hand writing.

561. SINGH, S. P., *High School Examination, Aptitude and Teachers' Estimate as Predictors of Achievement in Science at the Intermediate Level, Ph.D. Edu., Gor. U., 1975.*

The objectives of the study were : (i) to classify pupils seeking admission in intermediate class on the basis of their potentialities for science, (ii) to help increase efficiency of the individual for better understanding of science by guaging his specific aptitude according to the nature of the subject, and (iii) to determine predictive power of high school examination marks at intermediate level.

A battery was developed to predict success in achievement in science at intermediate level. The components of the battery were selected through job analysis and experts' opinion. Seven tests were included in the battery. These were : (i) manual dexterity, (ii) cause-effect relationship, (iii) spatial perception, (iv) reasoning, (v) observation, (vi) numerical ability, and (vii) memory. The tests were administered to a randomly selected sample of science students studying in intermediate classes in two educational administrative regions of Eastern Uttar Pradesh. The sample included students of either sexes from rural and urban schools. The average age of the sample was 15 plus years. Regression equation was derived for selection. The discriminant function was applied to classify the individual either to the biology group or to the mathematics group. Regression of scores was calculated so as to introduce standard procedure in the screening process. Pooling square technique was followed to see the relative correlation

of such predictors, viz., previous examination grade, teacher's estimate and personality inventories.

The important findings were : (i) the multiple R and composite reliability of the battery were found to be 0.49 and 0.85 respectively; and (ii) a regression equation to predict achievement in science was developed.

562. TEWARI, S. R., *Factorial Analysis of Areas of Attainments in the Science and Mathematics Courses at the High School Level, Ph.D. Edu., Gor. U., 1975.*

The study aimed at (i) exploring the psychological linkage among school subjects which required cognitive abilities for achievement and could be variously combined, and (ii) analying the correlations among various branches of science and mathematics to uncover the underlying functional unities in terms of which the grouping of subjects at the elective stages might be reconsidered with a side effect on the possibilities of related studies of sets of subjects which might enforce each other by reason of common elements which accounted for the achievement.

The following achievement tests were constructed and administered to assess the achievement of students in areas of attainments in the science and mathematics courses at the high school level, (i) arithmetic, (ii) algebra, (iii) geometry, (iv) physics theory, (v) physics practical, (vi) chemistry theory, (vii) chemistry practical, (viii) zoology theory, (ix) zoology practical, (x) botany theory, and (xi) botany practical. The tests were constructed after the tryout of a large number of items and item analysis. The reliability of the tests was computed by Spearman-Brown formula and Kuder-Richardson method. The skewness of distribution of the test scores was tested with the help of percentile formula. The validity of the tests was taken for granted as they were the achievement tests. Factor analysis was done by Thurstone's centroid method. Medland's method was followed for the estimation of communalities throughout the process of factor analysis. For the better approximation of communalities reiteration was done. Leyard Tucker's criterion and Burt's empirical formula for standard error of factor loadings together with the Vernon's recommendation were used to test the significance for further extraction of factors.

As a result of analysis five factors were extracted, namely, (i) Theoretical factor, (ii) Mathematical factor, (iii) Biological Science factor, (iv) Chemistry-Botany factor, and (v) General Science factor. The investigation revealed that curricular achievement align-

ed due not only to the innate abilities combining with aptitudes but also due to the imposition of complex functions by training and education upon simpler innate structures. The investigation brought out that algebra and geometry facilitated learning of various branches of science. The symbolic elements of algebra and figural elements of geometry were found helpful in developing numerical and spatial reasoning.

563. TLUANGA, L. N., Examination as a Mode of Management, Ph.D. Edu., Gau. U., 1974.

The major aim of the study was to present a systematic analysis of the prevailing system of essay type mass examination in Indian universities as a mode of measurement with special reference to the process of marking and the uncertainties in the examination marks owing to the variabilities of examiners' marking standards.

Critical investigation was made in respect of multiple marking and some other aspects of marking, such as, the theoretical soundness or otherwise of the practice of engaging a third examiner when two examiners differed beyond certain limit and the extent to which an examiner's impression of a script was carried over to the succeeding scripts. For studying multiple marking, a stratified random sample of 100 scripts in English from B.A. Part I Examination of Gauhati University was drawn. The scripts were marked independently by nineteen examiners scattered all over India. Error analysis was made in respect of constant error, deviation error and random error. Mark sheets were analysed statistically by computing central values and dispersion of marks. A four percent sample study of the matriculation results in 1963 was also made.

The following were some of the significant findings of the study : (i) strict examiners were more accurate than lenient examiners; (ii) marking standard was independent of marking spread; (iii) for the raw mark of an arbitrary examiner, the excess over pass mark should be about ten so that the script would pass on true marking; (iv) if a single examiner marked a pair of scripts, one could accept his order of merit with ninetyfive percent confidence if he put one script at least ten marks above the other; (v) if two examiners marked a pair of scripts, the order of merit could be accepted with ninetyfive percent confidence whenever the mark difference, after adding, was at least eleven; (vi) there was no justification for referring a script to a third examiner on the ground that the first two examiners gave divergent marks; (vii) the mean mark of two independent examiners was substantially more accurate than a single mark; (viii) out of three marks if one relied on the two in agreement rather than on the mean of all the three, the reliability of the assessment would decrease; (ix) from a study of the 'persistence effect' in marking, it was found that an examiner's judgment when marking a script was influenced by his impression of the preceeding script; the average error due to this effect and other sources of fluctuation in marking ranged between four to ten percent both above and below the mean.

*564. VARMA, P., Success-norms in Hindi Language for Classes VI and VII, Ph.D. Edu., Pat. U., 1977.

The important objectives of the study were : (i) to measure the success of students in the following written language abilities, viz., word meaning, language uses, grammatical form recognition, sentences, analogy of words, and comprehension; and (ii) to prepare tests to achieve the objective (i).

The sample consisted of only 50 students of classes VI and VII for preliminary tryout study, 100 students of classes VI and VII for tryout and item analysis, and 500 students of classes VI and VII for standardisation. The final test consisted of 20 word meaning items, 23 language use items, 15 grammatical form recognition items, 21 sentence items, 20 analogy of words items and 6 comprehension items.

The test-retest reliability was found to vary from 0.61 to 0.84 and split-half reliability corrected to full length was found to vary between 0.72 and 0.88. The validity was found against the annual examination scores. It was found to be 0.72 for class VI and 0.69 for class VII. Percentile norms were established.