THE UNIVERSITY OF HULL

AN ANALYSIS OF SOME ASPECTS OF THE RÔLE AND PERFORMANCE OF AN 'O'-LEVEL BIOLOGY EXAMINATION AS AN INSTRUMENT FOR CURRICULUM EVALUATION. A CASE STUDY OF THE 1983 'O'-LEVEL BIOLOGY EXAMINATIONS IN TANZANIA.

being a thesis submitted for the degree of

Doctor of Philosophy

In the University of Hull

by

Agnes Fellicia Elimankinga Njabili,

B Sc.(Educ) Hon.(Dar-es-Salaam), M.Ed. (UCLA).

August, 1985.

Summary of Thesis submitted for D.Phil. degree

by A. F. E. NJABILI

on

An Analysis of Some Aspects of the Role and Performance of an 'O'-level Biology Examination as an Instrument for Curriculum Evaluation. A Case Study of the 1983 'O'-level Biology Examinations in Tanzania

An analysis of the Tanzanian O-level biology curriculum and the 1983 O-level biology examinations has been carried out in order to identify problem areas.

The topics on which candidates performed worst were also among the topics that candidates and their teachers perceived as being difficult.

Item analysis showed a wide variation in the facility and discrimination indices across the items of the Theory and Practical examinations. Topic facility indices and ability facility indices were low.

The estimates of internal consistency reliability revealed that the reliabilities of both the Theory and Practical examinations were within acceptable ranges.

Importance of objectives as both course objectives and examination objectives was studied, and the abilities concerned with scientific method were accorded highest priority by both teachers and panel members responsible for the general oversight of the curriculum and examinations.

The classification of the questions according to the ability areas and topic areas tested by the examinations showed that all the ability areas and topic areas were examined, although to differing degrees.

The priority listing of the aims of Practical examinations in biology in Tanzania were similar to priority listings of practical aims reported elsewhere.

By factor analysis, it was revealed that the constructs underlying performance were related both to topic areas dealt with by the items, and to curriculum objectives tested. Some evidence was found that a 'Theoretical' mode of performance, a 'Practical' mode of performance, and a 'Continuous assessment' mode of performance constituted separate constructs.

A comparison of the performance by different sexes in the different ability areas tested by the examinations showed an overall girls' superiority. There was no indication of sex superiority with regard to performance in the different topic areas tested; or in performance in Continuous assessment. Both sexes performed better in Continuous assessment than in the Final examination.

Overall, the examinations seem to have been more suited for the selection of the candidates for entry into the limited places in higher educational institutions and for selected careers placement, than for evaluating the effectiveness of the course.

To:

My Mother	-	Ndeanshikoya Elimankinga Mushi
My late Father	-	Elimeleck Joshua Sandi
My Husband	-	Grant Wilfred Obel Njabili
My Sons	-	Baraka Obel Njabili
	-	Amani Elimeleck Njabili

for their love and suffering to see me through my educational career.

ACKNOWLEDGEMENT

I extend my most sincere thanks to everyone who contributed through different ways towards the completion of this study.

Foremost, I am thankful to the Government of Tanzania who sponsored me, and the Director of the Institute of Education, for granting me the study leave.

Equally foremost, I am deeply indebted to my Supervisor, Mr. C.R. Brown of the Science Education Centre, University of Hull, for his tireless supervision, readiness to help me at any time, and his most useful criticisms and suggestions which helped me shape the work.

I am grateful to Dr. W.J. Wilkinson, also of the Science Education Centre, University of Hull, for his suggestions on the appropriate statistics for data analysis, particularly his suggestions on procedures for factor analysis.

I am grateful to the different institutions in Tanzania that contributed to this work. Foremost in the list, is the Examinations Council, for allowing me to use the data from 1983 candidates' final examination scripts and continuous assessment scores. I am in particular most thankful to the staff of the Research and Evaluation Department of the Council for making sure that my needs were attended to.

My gratitude goes to Mr. E. Shitindi, of the Research and Evaluation Department of the Institute of Education, for helping tirelessly with the coding of the raw scores from the candidate scripts to coding sheets for computer analysis. Without his help this study would have taken much longer to complete.

I am indebted to: the students who completed the pupil questionnaire; the biology panel members and the biology teachers, who also completed questionnaires; and most of all, the 8 biology examiners from England and Wales whose classification of the questions and curriculum objectives constituted a neutral base from which I could make some statements about the examinations.

I am grateful to Professor V.A. McClelland, Director of the Institute of Education, University of Hull, for his assistance in attending to my financial correspondences. To the same effect I am grateful to Mrs. F.J. Owen and Dr. I.A. Pickering, of the Administration Department, University of Hull, for their endless efforts in making my financial position known to my sponsor.

I am most grateful to the staff of the Education Library and Brynmor Jones Library, University of Hull, who were so helpful to me whenever I needed materials from the library, for processing inter-library loans, and for other services rendered.

I am thankful to all the staff of the University Computer Centre for their assistance in various ways during the data analysis for this study.

I am grateful to Mrs. K. Atkinson for typing the instruments, and for being such a good friend for the whole time I stayed in Hull.

I am thankful to Mrs. E. Shearsmith for her endless efforts and time spent on typing the final work within the short time available.

I am indebted to Mr. E. Nkumbi and Mr. P. Masanja, my Tanzanian countrymen in Hull, for their moral support and all help rendered in the course of this study, particularly during the times when Obel and Eli were ill.

Finally, and most important, my gratitude goes to my husband, Grant and my sons Obel and Eli, for their love and perseverance to see me through this study.

TABLE OF CONTENTS

Page

Page

Page

List of Tables.

Page

Page

Page

Page

List of Figures

List of Appendices

ABSTRACT

The O-level biology syllabus (1976); Teachers' guide (1978); 1983 O-level examination papers; and responses from eight hundred and fifteen (815) 1983 O-level biology candidate scripts; five hundred and nineteen (519) student questionnaire responses; forty nine (49) biology teachers' questionnaire responses; eight (8) biology members questionnaire responses; and eight (8) biology examiners' classification of questions and curriculum objectives constituted the main data source for:-

1. a detailed examination item analysis.
2. identification of topics of difficulty of the curriculum by examination item analysis (actual difficulty), and by teachers' and students' perception of difficulty.
3. estimation of reliability of the examinations by an internal consistency method.
4. determination of teachers' and panel members' views of the importance of the curriculum objectives as course objectives, and as examination objectives.
5. determination of the content validity of the examinations.
6. determination of aims of Practical work in biology in Tanzania.
7. a study of the precise nature of the items of the Theory and Practical examinations.
8. determination of the construct validity of the Theory and Practical examinations; and an investigation of whether Theory, Practical and Continuous assessment constitute separate modes of performance.
9. comparison of performance by different sexes on the Theory, and Practical examinations and on Continuous assessment.
10. comparison of performance on the Final examination (FE) and Continuous assessment (CA).

The detailed item analysis showed a wide variation in the facility and discrimination indices across the items of the Theory and Practical examinations. The results of the item analysis were not found to be related in a systematic manner to the abilities tested by the items. The ability facility indices (AFI)

were low.

However, there was a relationship between item facilities and topic areas tested by the items. The topics on which the candidates performed worst were also among the topics of highest perceived difficulty. Some topics which were thought to be easy did not give rise to high performance in the examination; while some topics which were thought to be difficult did give rise to high performance in the examination.

The estimates of internal consistency reliability revealed that the reliabilities of both the Theory and Practical examinations, although capable of improvement, were found to be within acceptable ranges. The Practical examination yielded higher reliabilities than the Theory examination.

With regard to the importance of the objectives as course objectives and examination objectives, although some teachers' priorities differed from panel members' priorities, the abilities concerned with scientific method were accorded highest priority by both groups.

The examiners' classification of the questions according to the ability areas and topic areas tested by the examination items, showed that all the ability areas and topic areas were examined through the two examination papers, although to differing degrees. Therefore, while some of the objectives were accorded the weights which were thought to be desirable by teachers and panel members, some were underweighted. The classification of the questions into the six categories of Bloom's Taxonomy showed that the knowledge and comprehension categories appeared to have been covered more extensively than the higher cognitive categories.

The determination of the aims of Practical examinations in biology in Tanzania revealed a close agreement between teachers' and panel members views and were similar to priority listings of practical aims reported elsewhere.

The determination of the nature of the items of the Theory and Practical examinations showed that the constructs underlying performance were related both to topic areas dealt with by the items, and curriculum objectives tested.

In the determination of whether the components of each paper constituted separate constructs, and if performance in theoretical and practical biology examinations constituted separate constructs, it was shown by a multitrait-multimethod matrix approach that there was strength for divergent/discriminant validity of the components of the Theory and Practical examinations. The experimental investigations section of the Practical examination, and the project component of the CA showed the greatest divergency. By factor analysis it was revealed that 'Theoretical' mode of performance and 'Practical' mode of performance are separate constructs. In addition, a 'Continuous assessment' mode of performance was shown to be separate from the 'Theoretical' and 'Practical' modes of performance.

The comparison of performance between different sexes in different ability areas tested by the Theory and Practical examinations showed an overall girls' superiority. However, there was no indication of sex superiority with regard to performance in the different topic areas tested; or in performance in Continuous assessment.

The comparison between performance in FE and CA showed that performance in Continuous assessment was higher than performance in the Final Examination for overall sample,girls and boys samples similarly.

CHAPTER ONE

CURRICULUM AND EXAMINATIONS IN TANZANIA

1.10 Introduction.

Tanganyika, which was a British Territory, received its independence on the 9th December, 1961. The United Republic of Tanzania was formed in April, 1964 by the union of mainland Tanganyika with the islands of Zanzibar and Pemba.

Before 1961 the education system in Tanzania was modelled on the British system of education, and was designed for a highly developed society in which both primary and secondary schooling were available for all, and in which the skills required were predominantly clerical and industrial ('white-collar' skills).

There were three predominant features which characterized schools and school curriculum in Tanzania before 1961:

1. Schools were organized and run on a racial basis. There were separate schools for Whites, Indians and Africans. The poor schools (those with less equipment, which were understaffed and had poor class room facilities), were always allocated to the Africans.
2. There were very few schools, especially secondary schools; where they occurred they were built in or near urban centres.
3. The curriculum content was foreign. The curriculum materials, were mainly examination syllabuses prepared in Britain.

Likewise, the textbooks, and other teaching materials which were recommended for use by the examining body, The University of Cambridge Local Examinations Syndicate, were developed and written in Britain for use in British schools, e.g. those for English Literature and British Regional Geography.

There was no set curriculum to be followed, teaching and the selection of teaching materials were determined by the requirements of the examination syllabuses, regardless of relevance to the nation.

Immediately after independence in December 1961, changes were made in the education system. The first major change was the elimination of racial distinction within education. This was an easy task to accomplish as it only involved enrolling mixed races in all the schools.

The second change involved the expansion of educational facilities. Educational expansion was one of the priority areas in the First - Five - Year Development Plan of 1963 - 1968. Pronounced expansion took place at the Secondary Schools and Post - secondary school levels. Thus, by March 1967, the number of students attending secondary schools was 25,830. This was more than double those in attendance in 1961 when the figure was 11,832. (Nyerere, 1967).

The third change involved the modification of educational content. Efforts were made to introduce changes in existing syllabuses so that the content of education was more relevant to Tanzania. The effects of change in curriculum content were felt more in subjects like history, geography, and civics than in other subjects. This was achieved by the efforts of University Lecturers, and staff of the Institute of Education, and the Ministry of Education producing teaching materials for use by the teachers in the field. At that time, 1967, most of the new materials were circulated as pamphlets. The pamphlets were later modified and published as textbooks for the schools, for example, "History for East and Central Africa" (Education, Institute of, 1975). There was, very little change, if any, in science syllabuses.

1.20 Education for Self-Reliance (ESR).

In January, 1967 a policy document titled 'The Arusha Declaration' was issued by the ruling party, Tanganyika African National Union (TANU). The document gave a description of the kind of society that Tanzania intended to build. In summary, the document called for the creation of a socialist society based on four principles:

1. Equality and respect for human dignity.
2. Sharing of the resources which are produced by peoples' efforts.
3. Work by everyone and exploitation by none.
4. Struggle for Self-Reliance. (TANU, 1967).

The emergence of a national system of education was inevitable following the rapid political changes which were affecting the country. Education has been recognized as being one of the most effective tools for bringing about social change, as Owen puts it; 'What a child learns will influence how he looks at life around him. If a state tries to control how people see the world rather than leave them to their own interpretation, curriculum is a useful instrument of control. It can, early on, indoctrinate and prescribe'. (Owen, 1973, p.5.). He goes on to say 'To indoctrinate is not always undemocratic or, in one sense, wicked'. A policy document outlining the purposes of education in Tanzania was issued in March 1967 by the President of the United Republic. This policy document is known as 'Education for Self-Reliance' (ESR). The policy identifies the purposes of education in Tanzania as follows. To:

1. Foster the social goals of living together, and working together, for the common good.
2. Prepare the young people to play a dynamic and constructive part in the development of a society in which all members share fairly in the good or bad fortunes of the group.
3. Inculcate a sense of commitment and help the pupils to accept the values appropriate to the kind of future called for. (Nyerere, 1967 p.273).

This means that the education system in Tanzania must emphasize co-operative endeavour and responsibility to give service which goes with any special ability whether it be in manual skills, such as carpentry, in animal husbandry, or in academic pursuits; counteract the temptation to intellectual arrogance; and prepare young people for the work they will be called upon to do in the society which exists in Tanzania - a poor rural society with a predominantly agricultural economy.

ESR included an analysis of the Tanzanian education system in 1967, which revealed that:-

1. The education provided was still designed to meet the interests and needs of a small proportion of those who entered the school system. An academic tradition existed which aimed at preparing secondary school leavers for professional jobs, and academic life.

2. Those who went through the education system became divorced from the society for which the education was supposed to prepare them.

3. The education provided encouraged the thought that worthwhile knowledge was only acquired from books or from 'educated people' (those who have been through the formal education system).

4. The young, strong people, (the school population), were not involved in productive work. They were merely consumers. (Nyerere, 1967).

To alleviate the above situation, the policy for Education for Self-Reliance called for a change in the content of the curriculum; the organization of the schools and the entry age into primary schools.

This study will not deal further with the last two elements of change mentioned above. Curriculum change, and efforts made towards curriculum change in Tanzania form the focus of the rest of this chapter. The theme of curriculum development and curriculum evaluation is discussed further in Chapters Two and Three.

1.30 Curriculum Development and Curriculum Evaluation in Tanzania.

The translation of the purposes of education as laid down in ESR, into curriculum issues has been the concern of the education system since ESR was published.

For curriculum change to take place at national level, key decisions are necessary. Decisions need to be made about:-

1. The over-all educational aims which schools are to pursue.
2. General and more specific objectives of instruction.
3. The subjects, or other major areas of the curriculum.
4. The specific content to be covered in each selected area.
5. The choice of the types of learning experiences with which to implement the educational objectives.
6. How to evaluate what students are learning and the effectiveness of the curriculum in attaining the desired ends.

As mentioned, in Tanzania, the general educational purposes have been identified in the policy document titled 'Education for Self-Reliance'. The goals of education have been interpreted in the following general aims by the Ministry of Education:

1. To equip learners with knowledge, skills and attitudes for tackling personal and societal problems.
2. To prepare the young people for work in Tanzanian's predominantly agricultural society.
3. To equip learners with the values of freedom, unity and socialism (Education, Ministry of, 1980, p.3.).

All other decisions concerning curriculum are made with reference to the above aims.

1.31 Decision making bodies in Tanzania.

There are four bodies which are involved in making decisions affecting the secondary school curriculum: The Ministry of Education; The Institute of Education; The Examinations Council; and the schools themselves.

The Ministry of Education.

The education service in Tanzania is the responsibility of the Ministry of Education. It is the Ministry of Education which decides on:

i. The general objectives of instruction for the different levels, (primary, secondary and higher education).

ii. The subjects which constitute curricula at the different levels, and the over-all pattern of the curricula, for example, subjects to be taught at each level, and number of periods per subject per form.

A summary of these objectives of instruction and subjects of the different curricula can be found, for those interested, in 'Basic Facts About Education in Tanzania' (Education, Ministry of, 1980).

The rest of the decisions listed on page 5 are made by the Institute of Education, and/or The Examinations Council, in collaboration with teachers. Figure 1.01. shows the functional organization of the Ministry of Education and the relative organizational relationship of the Institute of Education, and the Examinations Council.

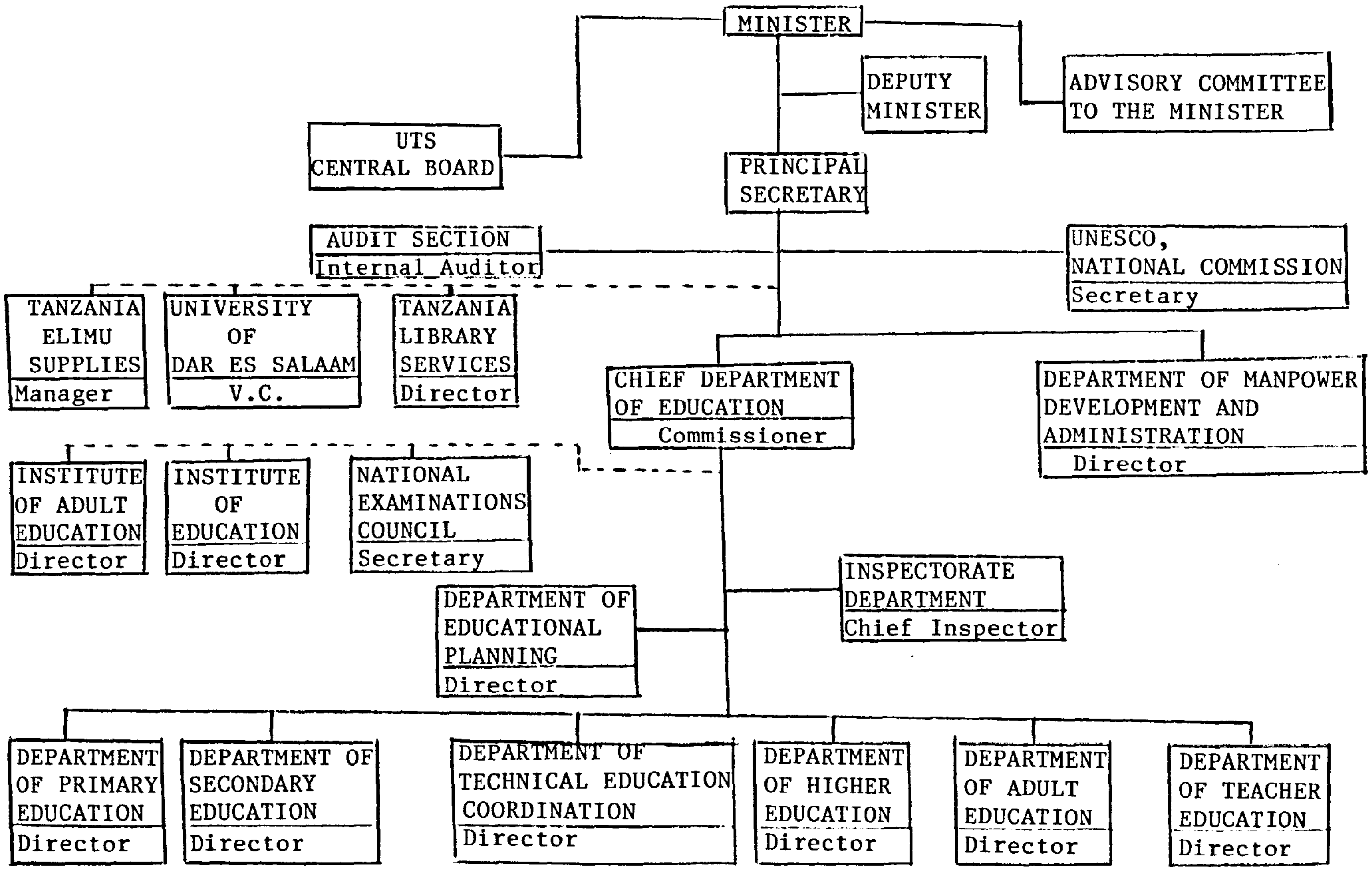
The Institute of Education.

The Institute of Education is a parastatal organization under the Ministry of Education. The powers to decide on most issues concerning the details of school curriculum have been vested upon the Institute of Education by an Act of Parliament of the United Republic of Tanzania (No.13 of 1975).

The Institute of Education was first established by the Council of the University of East Africa in 1964 as an educational body working under, and being run by Dar-es-Salaam University College. Among its many duties carried out between 1964 and 1975, the following specifically concerned curriculum development: Assisting the preparation of syllabuses for schools and teachers' colleges; promoting educational research and evaluation; disseminating information on new teaching methods, experiments and results of research works (Education, Institute of, 1982).

As already mentioned, much of the early curriculum development work focused on secondary education because the expansion of secondary education was greater in the First-Five-Year Development Plan than at the other levels (primary and post-secondary), and the available personnel working with the Institute were more familiar with the secondary school curriculum.

Fig. 1.01 ORGANIZATION CHART : MINISTRY OF NATIONAL EDUCATION - 1983



In the late 1960s, and early 1970s, there were some overlap and confusion in the roles of the Institute of Education and the Directorate of Curriculum Development in the Ministry of Education as regards issues of curriculum innovation and the development of curriculum materials. The main issue was on the question, 'Who should make the decisions on the content of the curriculum?' (Education, Institute of, 1982).

To eliminate the overlap, and the confusion, for example, in the issue of different syllabuses for the same subject at the same time, the Directorate of Curriculum Development was abolished, and the Institute of Education was established in 1975 by an Act of Parliament of the United Republic of Tanzania (No 13 of 1975) as a body independent from University of Dar-es-salaam, and answerable directly to the Ministry of Education.

The following powers are vested upon the Institute of Education by the Act:-

1. To assume responsibility for the development of educational programmes and to undertake the evaluation of courses of study and practices on the basis of national educational objectives.
2. To undertake analysis, review and revision of curriculum and syllabi.
3. To initiate, promote and supervise such changes in the syllabi and educational programmes as are necessary for the implementation of the national policy on education.
4. To conduct training programmes in such subjects associated with the development of curricula and syllabi.
5. To collect and make available to the government and other public authorities, information on methods of teaching, content of courses of study and current technological developments in education.
6. To give advice and make recommendations on such matters relating to curriculum reform.
(Education, Act of the Institute of, 1975).

To that effect, the Institute of Education is responsible for undertaking the following roles:

i. Selection of the subjects of the curriculum or major areas in terms of specific content to be included in each area.

ii. Choice/suggestion of types of learning experiences with which to implement both the content and the objectives.

iii. Together with the Examinations Council, deciding on how to evaluate what students are learning and the effectiveness of the curriculum in attaining the desired ends.

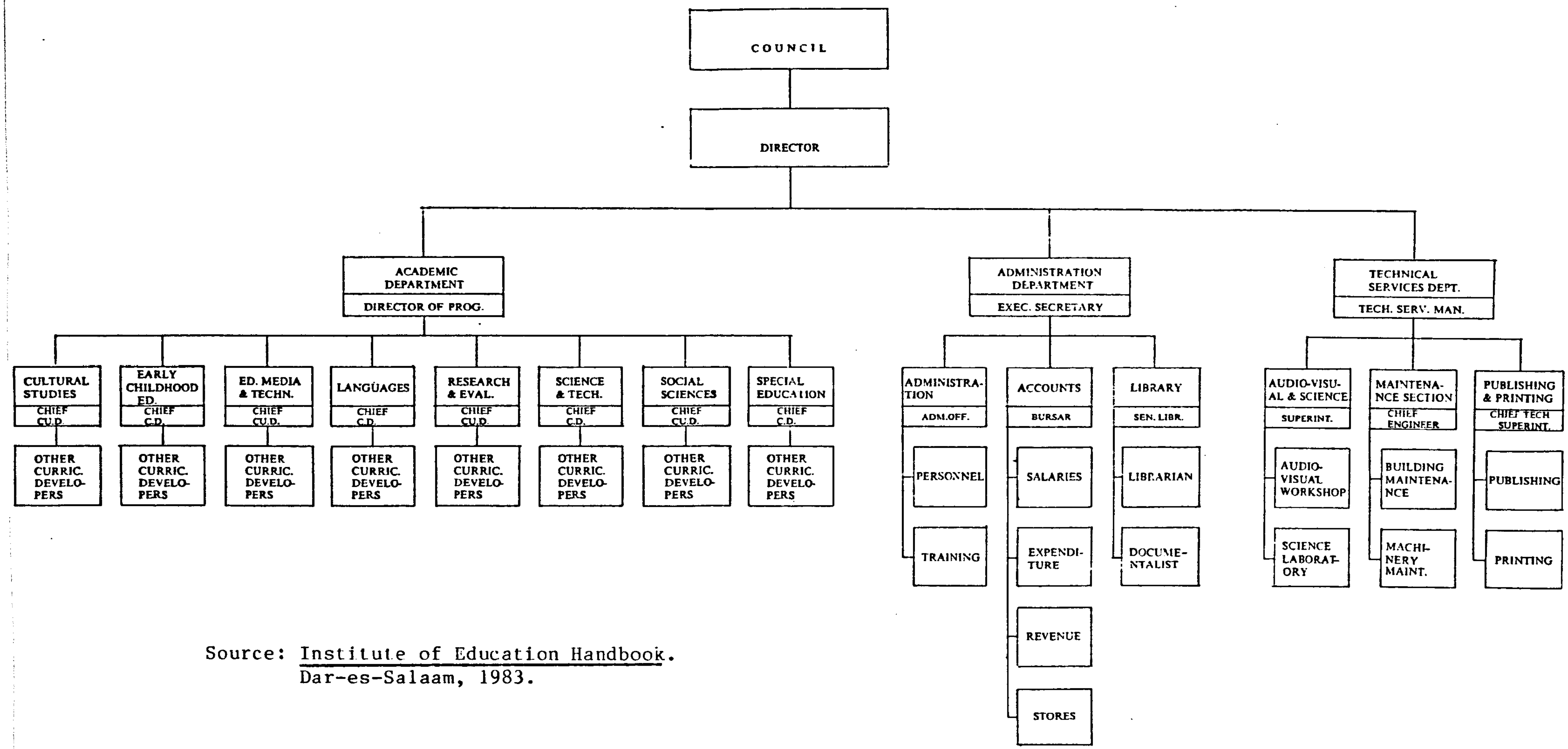
iv. Together with the Ministry of Education, deciding on what the overall pattern of the curriculum is to be.

The Institute of Education operates a series of subject panels which are therefore its functional units. Subject panels were first formed immediately after the announcement of ESR. They comprise teachers, Ministry of Education representatives, and subject specialists from the University of Dar-es-salaam. The minimum panel membership is fifteen (15) and the maximum is twenty six (26) members.

Subject panel members meet often to discuss, devise and/or revise syllabuses, textbooks, teachers' guides and other teaching materials, and to discuss matters pertaining to the teaching-learning situations in the schools. During such meetings, they also discuss and devise teaching strategies, and make suggestions on areas where research and evaluation need to be conducted. The Institute has a full-time curriculum developer for each subject who acts as a secretary to all such meetings. In addition to being a full-time member of the panel, the secretary's main duty is to co-ordinate all activities of the panel and to organize the work done by the panel.

Most panel work is done through workshops and seminars. It is worth noting that not all the work of the panel is done by every panel member. Individual members can be assigned to prepare specific drafts of materials which may be discussed by small sub-groups of the panel, or by the entire panel, depending on the nature of the materials. Policy decisions, however, are taken by the entire panel. Fig. 1.02 elaborates the functional organization of the Institute of Education.

Fig. 1.02. **APPROVED ORGANIZATION CHART OF THE INSTITUTE OF EDUCATION**



Source: Institute of Education Handbook.
Dar-es-Salaam, 1983.

The curriculum development work carried out by the Institute of Education is thus an integral part of the implementation of the national educational policy-ESR.

The curriculum documents produced, such as syllabuses, teachers' guides, and textbooks are sent to individual schools for use. That the curriculum sent to schools is followed, is checked by school inspectors who are directly answerable to the Ministry of Education through the Directorate of Inspectorate. Schools can expect to be inspected at least once a year. That prescribed curriculum is followed is bound to occur because the Examinations Council bases its examinations on these approved curricula sent to schools.

Compared to England and Wales, the control of curriculum in Tanzania can be said to be centralized. In England and Wales, the responsibility for the curriculum has been vested in the Local Education Authorities, (LEA), and in the Governors and Managers of Voluntary Schools (Education Act, 1944 - Section 23). Though it can be argued that the system is centralized at LEA level, in practice, however, decisions about the curriculum are taken by teaching staff at the level of the individual schools. As in Tanzania, the secondary school curriculum is controlled by a system of external and externally moderated examinations. The difference between the examination system in England and Wales and that of Tanzania is that in England and Wales the actual examinations are administered by autonomous examination bodies, while in Tanzania, they are administered by one body, The National Examinations Council.

The Examinations Council.

Before independence, (1961), University of Cambridge Local Examinations Syndicate was the examining body for all school examinations in Tanzania. The Syndicate continued in this role after independence, up to 1967.

In November 1967, the East African Examinations Council was established to take over the organization of the School Certificate and Higher School Certificate examinations for Uganda, Kenya and Tanzania. However, this new arrangement did not last long because each country possessed a distinct political ideology and seemed to demand that the education provided, and therefore the examinations, should be more relevant to its own social and political aspirations (Examinations Council, 1981).

In April 1971, the Ministry of Education expressed to the Cabinet the wish to conduct its own O-level and A-level Secondary School Examinations for the country. In addition to the point made in the previous paragraph four further reasons were given:-

1. That so far, Tanzania had managed to conduct successfully its own Primary School Leaving Examinations (PSLE).
2. That the University of Dar-es-Salaam, under Tanzanian leadership was running and managing its own examinations.
3. That it was only at the secondary school level that the examinations were externally set and supervised.
4. That the costs of running the externally set and supervised examinations were very high.

(Examinations Council, 1981).

The Ministry of Education, were able to draw on the expertise of personnel who had served on the East African Examinations Council Board in order to set examinations for the Tanzanian candidates who had registered for the 1971 Certificate of Secondary Education, and the Higher School Certificate Examinations. Thus the first nationally set examinations were taken in November 1971.

Later, in December 1973, the National Examinations Council of Tanzania was established by an Act of Parliament (No.21 of 1973). Among the duties, vested upon the Council by the Act is the duty 'to conduct examinations for, and grant diplomas, certificates, and other awards of the Council'.

Therefore, the Examinations Council is responsible for setting and marking the examinations and awarding certificates for primary schools, secondary schools, and colleges of education. The setting, moderation, invigilation and marking of all O-level and A-level examinations is determined and organized centrally by the Examinations Council. Fig. 1.03 elaborates the functional organization of the Examinations Council. It is this organization which is responsible for the O-level biology examinations which form the substance of this research.

From the information given previously it can be seen that there is a division of educational labour in Tanzania, in which the educational philosopher, (The Ministry of Education), defines the goals of education; the curriculum maker, (The Institute of Education), devises methods for attaining those goals; and the test constructor, (The Examinations Council), devises instruments, (the examinations), to measure the extent to which the goals have been attained.

One of the potential problems in such an arrangement is that of clear communication of ideas between the different parties. In Tanzania successful liaison is achieved, to some degree, by having each of the three institutions helping each other achieve its functions through participation in activities organized by these institutions separately and/or jointly, for example, workshops, seminars and examinations'marking. It may be that effective communication of this sort occurs in a centralized system of education, some of the characteristics of which are discussed in Chapter Two.

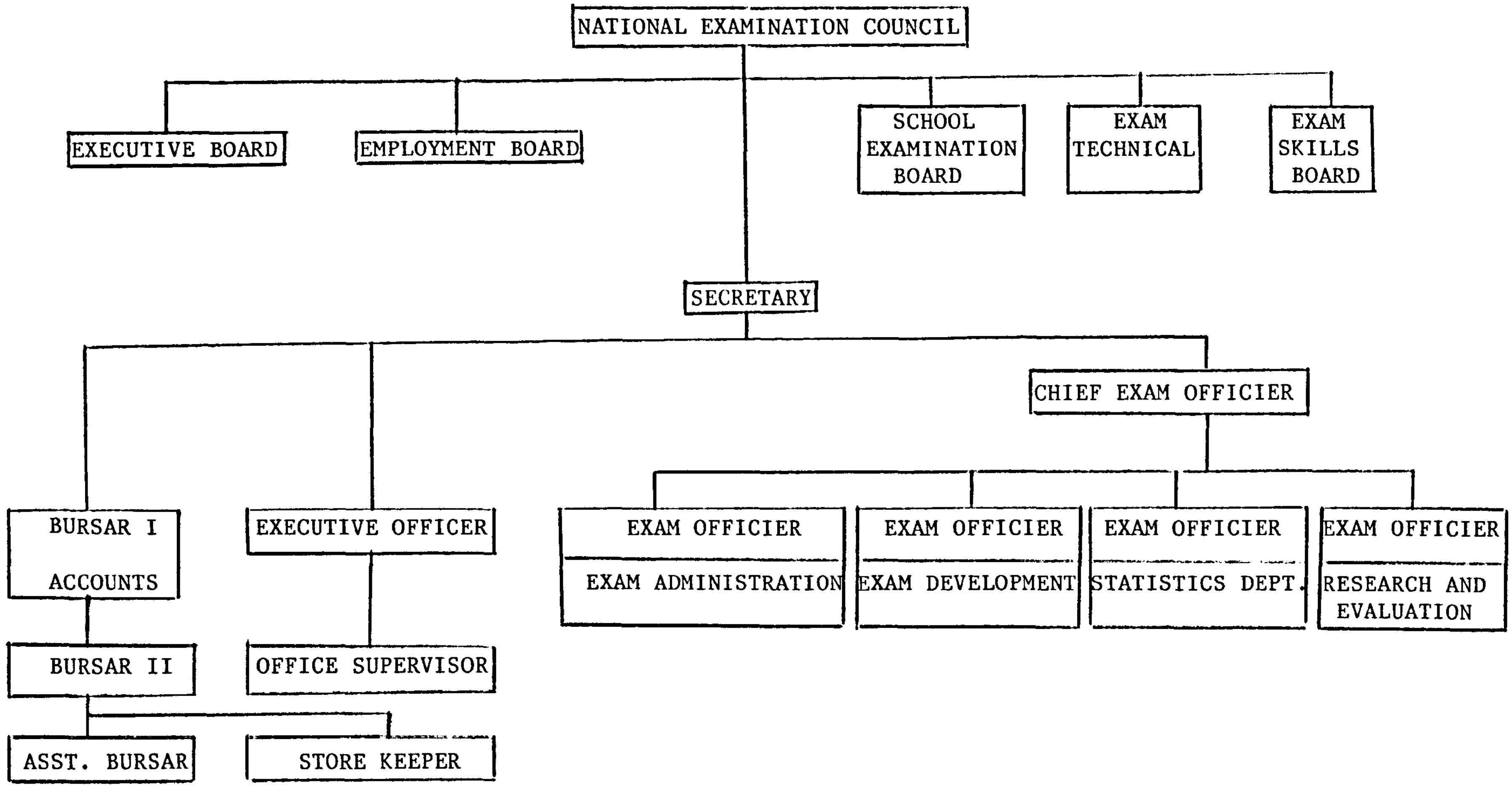
Teacher's contributions to Curriculum Development and Assessment of Student Achievement.

Teachers play a major role in curriculum development and assessment of student achievement. They do so individually, and in collaboration with the Institute of Education and the Examinations Council.

Individually, at school level, teachers decide on specific instructional objectives, often by modifying the objectives of instruction outlined in teachers' guides, to suit specific classroom situations.

In collaboration with the Institute of Education, they decide the content

FIGURE 1.03 NATIONAL EXAMINATION COUNCIL OF TANZANIA : ADMINISTRATION STRUCTURE.



Source: National Examination Council of Tanzania, Ten Years of National Examinations, 1971-1981, Dec. 1982.

of the curricula. They do so through those serving on the subject panels, and through panel activities, for example, workshops and seminars.

Through the panels' workshops and seminars, they make suggestions on the syllabuses and other teaching materials being used, and initiate changes when and where need be. They help in the preparation of teaching and learning materials and give suggestions on possible approaches and methodology to be followed while trying to implement the demands of the different curricula.

Again individually, and in collaboration with both the Examinations Council, and the Institute of Education, teachers help to evaluate what students are learning, and hence the effectiveness of the curriculum in attaining the desired ends.

Each teacher participates in the Continuous Assessment Scheme (CA). The continuous assessment scores awarded to each student in the different subjects contribute 50% to the final School Certificate Examination score in each subject.

Again with regard to assessment of student achievement, with the supervision of the Examinations Council, selected teachers construct, moderate and/or mark examinations at 'O' and 'A' level.

1.40 Changes in Science Teaching in Tanzania since 1961.

In the early 1960s, the teaching of science was undergoing major changes world wide. For example, in the United States of America, the following bodies were formed and funded major science curriculum development projects: The National Science Foundation (NSF), The Physical Science Study Committee (PSSC), and the Biological Science Curriculum Study (BSCS). In the United Kingdom, The Nuffield Foundation (NF); and in Australia the Nuclear Research Foundation (NRF) were formed. Most of these curriculum development initiatives involved teaching science through a practical approach with an emphasis on enquiry/discovery or problem-solving by the pupil (Schwab, 1963, Kelly, 1971).

Within the field of school biology, perhaps the best known curriculum development projects were those of the Nuffield Foundation and BSCS group. These materials soon spread to other parts of the world. For example, in Israel, the BSCS materials were modified to suit local needs and were re-named the BSCS-Israel version. Likewise in Australia the BSCS materials were adopted with modifications. The Australian adaption of BSCS is called 'Biological Science: The Web of Life", abbreviated 'The WEB' (Lucas, 1980). The Nuffield materials were also adopted by some countries but with modifications. For example, both the BSCS and Nuffield materials were used in Tanzania to develop the School Science Project (SSP) materials described in Section 1.41.

As pointed out in the examples above, most of the initiatives advocating a different approach towards science teaching occurred in the developed countries. In order that these initiatives could spread and acquire an international coverage, UNESCO set up an internal action unit to bring the benefits of these new developments to the less advanced countries (UNESCO, 1966). A programme was started by which visiting lecturers (scientists and teachers) from the developed countries could visit the developing countries to enable local personnel to become acquainted with innovations in science education which had occurred. The programme which was drawn by UNESCO for extending Science and mathematics activities was known as "UNESCO PILOT PROJECTS".

In addition, the United Nations body made proposals to set up a complementary international fellowship programme to permit students and scholars from one country to study the new approaches in science education in other countries. Also an international science teaching centre for research and development of new approaches, materials, and media of science teaching was proposed. (UNESCO, 1966).

Following these proposals, UNESCO drew up a programme to extend science education activities to developing countries. Pilot projects for Physics, Chemistry and Mathematics were established in Latin America, Asia and in some

Arab States respectively. In Africa, specifically Ghana, Biology pilot projects were started in 1967/68. (Meena, 1978).

Establishment of Science Panels in Tanzania:

In response to ESR; the UNESCO recommendations of 1966; and to follow suit in world trends, a joint science subjects specialists advisory committee, (the committee was first formed in 1965), to the Ministry of Education conference was called immediately after the announcement of ESR. At the conference, objectives for teaching science, and general lines of approach for teaching science, (Biology, Chemistry and Physics), in Tanzania, were laid down.

Among the recommendations made by the joint meeting were the following:-

1. In common with reform movements in science teaching in the world, Tanzanian courses should concentrate on processes, principles, and techniques of experimentation rather than the accumulation of factual knowledge which pupils soon forget after leaving school.
2. There was need for a change in the aims and forms of the examination. The examination called for was one which would test understanding rather than rote learning: the major principles and ideas rather than facts only and minor details.
3. The proposed changes were to be implemented gradually, beginning with the first forms of secondary schools.
4. Separate courses for Physics, Chemistry and Biology were to be developed.
5. Works or programmes done elsewhere in the science curriculum such as PSSC, Nuffield Science, BSCS, and any others were to be adopted freely whenever some of the ideas and principles given in such programmes were thought relevant to the needs of Tanzania. (Education, Institute of, 1967).

The proposed project for the development and improvement of science teaching in Tanzania was later named the School Science Project (SSP), with separate initiatives in the three sciences which became known as SSP Chemistry, SSP Physics, and SSP Biology.

In biology, the proposed new approaches were to bear in mind that biology in Tanzania should:

i. Provide a useful part of the education for every Form Four leaver not only those, (about 20%), who could proceed to Form Five, but for all students most of whom were going to be the future adults of Tanzania.

ii. Be taught in such a manner that pupils would learn to think, to solve problems in order to cope with the changes in their environment and to continue to learn when they had completed the course. In order to make this possible biology must be made interesting to the pupils so as to inculcate the right attitude from the beginning.

iii. Be made as modern a course as any secondary school biology course taught anywhere else in the world. But care should be taken that as far as possible, the course was geared to local examples.

iv. Be designed appropriately for a developing country with a largely agricultural economy, and its major concerns such as hygiene, nutrition, and industrial problems, should be included in the curriculum. (Meena,1978).

Each of the above points was to be taken into consideration during the development of course objectives and course materials for SSP biology. At that time, the main course material for biology, to be referred to as the 'Traditional Biology'in this chapter, was the University of Cambridge Local Examinations Syndicate Syllabus for Overseas Centres.

Meena (1978), has pointed out three major deficiencies of this syllabus (the Traditional syllabus). The examination course did not encourage a learning-by-discovery approach. Secondly, the curriculum did little to support agriculture, or the application of biological knowledge which was important not only to pupils' future careers, but also to the community in which the pupils were going to live. Instead, the syllabus prepared them for passing the examination for entry to Form Five (Higher School Certificate Programmes) only. Thirdly, the textbooks which were used had usually been written by foreigners for their home students and as such the language and teaching-learning resources were unfamiliar to Tanzanian pupils.

1.41 SSP Biology.

As pointed out in the previous section, the Biology panel was formed in 1965. Members of the panel included University Lecturers attached to the Institute of Education, teachers and tutors from schools and colleges, and representatives from the Ministry of Education.

One of the early tasks of the panel was reviewing the Traditional syllabus in order to make changes in line with the proposed approaches mentioned on page 18. The second task was to develop materials for SSP Biology. In the development of SSP Biology, the first two years were mainly utilized in planning and writing draft materials. Draft materials were written for teachers' and pupils' texts. The teachers' texts showed how the course might be taught, and the pupils' texts guided the pupils towards learning the course, mostly through experimentation.

The writing workshops and the panel, bearing in mind the joint science panel recommendations, and the purposes of education in Tanzania as pointed out in ESR, conceived the objectives of the new biology course as follows:-

1. The course should prepare all pupils for life in the village as well as some of them for work beyond this level (that is beyond O-level).
2. It should emphasize, whenever possible, the biological principles and problems relating to agriculture, nutrition, health and development and conservation of natural resources.
3. The course should encourage practical projects and investigations. Theories, concepts, principles and definitions should be developed as a result of practical work or investigations rather than the use of practicals to support or verify them.
4. The investigations in the course should challenge the pupils to think

critically. Pupils should be involved in observing, measuring, recording, classifying objects or events, inferring and formulating ideas.

5. The various projects, and/or investigations included in the course, should be performed on a co-operative basis - that is, in groups so as to encourage critical discussions.

6. The course should, as far as possible, incorporate local examples and illustrations which are familiar to pupils of Tanzania or East Africa in general. However, the materials produced by other curricula developed in other parts of the world should be borrowed freely but with the reservations that anything borrowed must be adapted to fit local needs. (Meena, 1978; Education, Institute of, 1967).

The First Year (Form One) teachers' and pupils' texts were first put to trial schools in January 1968. Initially, 15 secondary schools were involved in the trials with approximately 30 classes of 36-40 pupils in each. In 1969, the number of schools increased to 26. By the end of 1971, there were Four Teachers' texts, and Four Pupils' texts in use in the trial schools for Forms I-IV.

The pupils' texts served a dual purpose: first as a laboratory manual, and secondly as a class text. Thus each pupils' book contained practical work, investigations, problems, questions, and textual materials. (Education, Institute of, 1967, 1968, 1969).

The teachers' books(or guides), accompanying each pupils' book specified the objectives of each chapter and gave information to help the teachers in the organization of lessons. Most answers to problems and questions posed in the pupils' texts were also given. Finally, a list of apparatus and other materials to be used for various investigations by the pupils, individually, in a group, or by the teacher for demonstration were also given in the teachers' books. (Education, Institute of, 1967; 1968, 1969).

Teacher Participation in SSP Biology:

The pupils' texts and teachers' guides for SSP Biology were planned first by a small group of panel members who laid down the outlines. These outlines were then brought to a workshop where a larger group of teachers, selected from schools and colleges discussed the outlines, and wrote the draft materials. The majority of the participating teachers were selected from the trial schools.

As in the Nuffield trial materials in the UK, (Kelly, 1971b), teacher participation was highly valued. The participation of teachers in Tanzania was important for three reasons:-
First, their experience was essential as they were in the best position to know the conceptual level of the materials being written and the suitable local examples to be used. Secondly, by their participation in the writing workshops, they would, at the same time, become orientated to the new materials, thus saving time and money for running an in-service training course for them. Thirdly, they would naturally teach the course with more confidence because they would feel that they were part of the entire exercise. (Meena, 1978).

Measurement of Pupils' Achievement in SSP Biology.

Classroom exercises, tests and term examinations were used to assess pupils' progress in SSP Biology. The tests and term examinations consisted of different kinds of items, such as multiple choice, matching, true-false, short-answer and essays. The end-of-course achievement was assessed through an O-level SSP Biology examination taken after four years.

The first O-level SSP Biology examination, was sat in November 1971 by the pupils who started the course in January 1968. The Biology panel, and Examinations Authority, (the newly formed examination body for Tanzania), identified and concentrated on four objectives for testing achievement in the SSP Biology programme, (Meena, 1978):

1. Ability to understand biological principles.
2. Ability to communicate information scientifically.
3. Ability to understand technical and practical skills.
4. Ability to relate principles and skills to community and other situations.

The end of course examination consisted of two written papers. Paper one, (2½ hrs.), consisted of a mixture of question types similar to those mentioned in paragraph one of this section. All the questions were compulsory.

Paper two, (1½ hrs.) consisted of 'practical' theory questions. That is, descriptive questions based on practical work done during the four years course. A practical examination was discouraged for one main reason: that specimens for practicals, and the condition under which practical tests would have been conducted, were not uniform throughout the country, and would produce inconsistency in pupil responses, and hence pupil performance (Meena 1978).

The results of the first O-level SSP Biology examination were reported as being satisfactory (Meena, 1978) and the examination appeared to fulfil the following design criteria,

i. of measuring attainment in a way suitable for selecting students for entry into Form V (Higher School Certificate Programme), and
ii. of reflecting the work of the course not only in terms of the subject matter covered, but also the involvement of students in activities similar to those undertaken in the course.

The life Span of SSP Biology.

The School Science Project for Tanzania turned out to be a short-term enterprise, for SSP Biology, (SSP Chemistry, and SSP Physics similarly), started being phased out of schools in 1971. The last SSP examinations were taken in November 1973, (Examinations Council, 1981), and by 1974, SSP courses had completely disappeared.

The present author could not find any literature that gives reasons for the disappearance of the SSP course, but it seems likely that SSP Biology was terminated before a full evaluation of it was done. If an evaluation had been conducted which had included difficulties experienced during its implementation, valuable comparisons with courses which were developed later, and with the Traditional Biology course which was being followed in the non experimental schools at that same time, might have been made.

However, there were open criticisms about SSP courses brought up at various teachers' meetings between 1972 and 1974, although there is no proof that these complaints constitute the reasons for the disappearance of the course. For example, at the Heads of Secondary Schools Annual Conference in 1972 and 1973, (Education, Ministry of, 1972, 1973,), those Heads of schools having Higher School Certificate Programmes, (A-levels), reported that they had often received complaints from science teachers teaching A-level classes that, students who had taken SSP courses were not coping with the demands of the A-level courses.

A problem seems to have occurred because, unlike in the development of Nuffield projects, there had not been A-level courses developed for SSP Projects. Pupils who had intended to pursue A-level courses and had passed their SSP O-level Examinations, had therefore been placed in A-level classes following Traditional syllabuses in which the teaching approach differed markedly from that of the SSP O-level course.

Complaints also emerged in 1972, at the Science Teachers' Association of the United Republic of Tanzania, (STAURT). Most of the complaints centred on teaching difficulties experienced while implementing the course demands of SSP courses. Some of the issues raised were as follows:-

1. Some topics were referred to as being difficult, too long and too detailed, for example, ecology and classification .
2. There were general complaints that the course had too many investigations that demanded experimental procedures, and that most of the

investigations were often 'unsuccessful.'

3. Because some of the experiments were long term it was necessary for more than one investigation to be in progress at the same time. This caused problems. Often students failed to make records of observations of the experiments in progress because the experiments were boring and/or students had lost interest. The sequential nature of the topics also meant that each topic should have been completed before proceeding to the next.

4. Teachers also complained of shortage of equipment. SSP Schools had not been provided with extra science equipment to meet the increased equipment demands of the courses. Borrowing equipment from other schools had proved necessary with all the attendant problems of late returns or losses in a situation where funds were short and replacements were not readily available locally.

5. Failure to complete parts of the course, especially experiments was also cited as a problem. Failure was explained as being due to the heavy demands of the course, where other heavy demands were already made, e.g. in trying to implement the non-academic elements of ESR (STUART, 1972).

1.50 <u>The Traditional Biology Course (1967-1983)</u>.

The SSP Biology course has been described in some depth because, the author feels that although the course has been terminated, it is to date, the only real innovation in biology teaching which has been attempted in the country; and it has influenced the design of the traditional curriculum and the style of examining.

Those schools which did not follow the SSP courses followed the so called Traditional courses in biology, chemistry and physics. The research undertaken in this study involves some aspects of the Traditional Biology course, and its terminal assessment procedures.

While the SSP Biology group of the Biology panel were concerned with the development of the SSP Biology course, the rest of the panel members were involved with the Traditional Biology course based on the Cambridge syllabus.

A number of changes have taken place in the Traditional biology syllabus between 1967 and 1983 in the teaching content; the methods of teaching and in classroom organization. The changes which have taken place in the teaching content have mostly been 'piece meal', involving subtraction of topics, additions, and re-organization of topic sequence. Group work has been the main teaching approach.

Examples of topics which have been added to the syllabus, (the syllabus still remains the main curriculum material), include:

Soil and Soil Conservation with emphasis placed on Agriculture;

Parasites and diseases of man with emphasis on the major endemic diseases in Tanzania, with emphasis on causes and control;

Nutrition with emphasis on food preservation and storage (the need for and the methods of); Genetics with emphasis given to the use of genetics for better animal and plant breeding.

To support the syllabus, a 'Teachers' guide to O-level biology syllabus' was issued in 1978. Other curriculum materials include recommended textbooks, most of them still prepared outside the country, but including more tropical examples, and carrying the label 'tropical edition', one suspects for the sake of marketability. Some of the recommended textbooks have been listed in the Teachers' Guide (See Appendix II).

Again, as far as the author is aware the changes in syllabus have taken place without any research having been done to assess the extent to which the previous changes have taken root in the classroom in the practical sense of effecting fundamental changes in what teachers and pupils actually do.

Measurement of student achievement in Traditional Biology.

There have, however, been considerable changes in the way measurement of student achievement in traditional biology has been carried out in Tanzania, between 1961 and 1983. There have been changes in the structure of examinations questions, examination format, and in the methods of assessment employed.

At the time of independence, (1961), student achievement at O-level was measured by examinations run and conducted by the University of Cambridge Local Examinations Syndicate. This continued to be the examining body up to 1967. Until 1967, assessment of student achievement in biology was through three examination papers, two theory plus one practical.

One of the theory, (2½ hours long), consisted of a number of short-answer questions based on all parts of the syllabus. All questions on this paper were compulsory.

The second theory paper, (1½ hours long), consisted of nine traditional-'essay type' questions. Candidates were required to answer five of the nine questions.

The practical paper, (1 hour long), with the number of questions varying between three and four, required candidates to make investigations on specimens provided and to record observations.

Between 1968 and 1970, the examining body changed due to the establishment of The East African Examinations Council and a number of changes have occurred in the examinations at 'O' level. The number of papers was reduced from three to two.

Paper one, (2½ hours long), consisted of nine questions. Candidates were required to answer question one, (a mixture of objective and short-answer items), and any four of the remaining questions, (traditional-essay type questions).

Paper two, (1½ hours long), was a practical paper on which there was no choice of questions.

In 1971, when Tanzania, through the Ministry of Education, started setting and marking its own examinations, O-level biology examinations were changed again. The number of papers remained the same. The format for paper one, (2½ hours long), changed in such a way that it consisted of three sections: Section I with objective type questions; Section II with short-answer type questions; and Section III with five essay type questions. Candidates were required to answer all questions in Sections I and II, and any two out of the five traditional-essay type questions in Section III.

Secondly, the time allowed for paper two, (the practical examination), was increased from 1½ hours to 2 hours. The increase in time was deemed necessary to cope up with the paper's demands. There were more questions and specimens for the candidates to study and to record observations.

The above format and structure continued until 1976. By far the greatest change in the examination followed the 1967 ESR policy document which had called for a change in national assessment procedures. The document points out that examinations 'as a general rule assess a person's ability to learn facts and present them on demand within a time period. They do not always succeed in assessing a power to reason, and they certainly do not assess character or willingness to serve'. (Nyerere, 1967, p.282).

In response to the ESR call a 'Continuous Assessment' scheme (CA) was introduced in 1976. In Tanzania, CA has 2 dimensions, an academic, subject specific one, and a subject independent element-character assessment. Continuous assessment is therefore, carried out for two main reasons: from the academic point of view, it is aimed at minimizing the disadvantage that a single examination is not able to sample adequately questions from all the curriculum topics within a subject although it may attempt to sample all or nearly all the objectives of the particular curriculum. Secondly, CA is carried out to cater for character assessment.

Character assessment score is a single letter grade (A, B, C, D or E), given to each student for each school year using specified criteria. The details of the criteria for character assessment are not within the scope of this study because the CA character assessment is independent of the CA academic assessment.

As in all other subjects, a CA score is included in the assessment of achievement in O-level biology. Since then, the format of the examination papers have also changed so that, currently there are two examination papers, a 3 hours theory paper consisting of three sections: Section I with three traditional-essay type questions; Section II with objective type questions (multiple choice, matching, and true-false); and Section III with short-answer type questions. Candidates are required to answer one out of the three essay questions of Section I, and all the questions in Sections II and III.

Paper two, a 2 hours practical paper, requires the candidates to carry out experiments and observations on given specimens. All questions are compulsory.

The present examinations and CA are discussed in Chapters IV onwards.

1.60 Evaluation of Science Curriculum in Tanzania.

Why did SSP Biology Fail?

A personal view of why SSP Biology, (SSP Chemistry, and SSP Physics alike), failed leads to the thinking that SSP courses for Tanzania ended as they started - without scientific research. The project may have been too ambitious and was initiated without adequate preparation. Thus, although the science panel seems to have had a reasonably clear idea of the nature of the problem (the curriculum reform needed) for which a solution was sought, there was no research done to enable the team to comprehend the nature of the innovation and evaluation of its potential usefulness; and how they were to mobilize the resources needed to meet the demands of the innovation.

It is important to note that there should be little room for assumption in curriculum innovations, because the management for change is neither clear nor self-contained. There are no magic rules and there are no formulae which are all-encompassing. The problems which confront education in Tanzania are diverse, not only in their nature and quality, but also in their points of origin. Therefore, we cannot simply regard the management of innovations as a matter of looking in one direction, of totally accepting global patterns of change in science education, or of regarding problems of implementation as falling neatly within one category.

Evaluation of Traditional Science Programmes.

Little research has been reported on the evaluation of the curriculum and examinations in biology or in other science subjects in Tanzania. The author, could find no published records of evaluation studies undertaken in physics and chemistry.

However, in Secondary Mathematics, Seka (1981), determined the areas of difficulty within the O-level Mathematics curriculum. Data were collected through achievement tests, opinion questionnaires completed by Mathematics teachers and pupils of different levels; and interviews conducted to complement the opinion questionnaires. Although the findings of the above study were used as a basis for modifying the Mathematics curriculum, the study failed to look into, and point out the contribution that the O-level Mathematics examination made towards the evaluation of student achievement at the end of the four year course. As in other subjects, it is the performance on the O-level examinations which have been used to criticize the curriculum, and/or the examinations (TANU, 1974, Kitosi, 1982).

Research studies conducted on the biology curriculum and examinations vary in depth and coverage. In an analysis to determine problems pertaining to the teaching of O-level biology using the 1976 syllabus, an opinion questionnaire was circulated to biology teachers by the Biology Panel of the

Institute of Education (Njabili, 1978). The data which were analysed came from 43 respondents.

A number of areas of concern were reported. Certain topics were identified as difficult to teach and to produce learning difficulties for pupils. These included Genetics, Evolution, Ecology and Classification. The specific areas of difficulty within these broad topic areas were not identified. The topics were thought to be too difficult for O-level and it was suggested by the teachers that these topics be included in the A-level syllabus only. Ironically, these are the topics covering most of the principles of applied biology. The report points out the need expressed by teachers for in-service training, to learn how to teach those topics successfully. Some teachers felt under-qualified to teach those topics. Concern was also expressed by a small minority about lack of clarity of the objectives.

Other issues pointed out in the report are the heavy teaching load, the poor academic qualifications of the teachers, and lack of textbooks and reference books.

Meena (1978) has undertaken a historical survey of the trends in teaching biology in Tanzania, and has pointed out the emphasis teachers put on the use of a theoretical approach to teach the subject. He further points out that, although some teachers try to include a practical approach as much as they can, their efforts are often limited by the great shortage of basic laboratory equipment, a feature he observed in most of the school laboratories he visited. An analysis of tests and examinations showed that the classroom tests and examinations employed by teachers tended to test mainly the knowledge and comprehension objectives of the syllabus.

In a study designed to determine the extent to which the final biology examination and continuous assessment are valid measures of objectives of O-level biology, Kitosi (1982), collected data by interview and questionnaires administered to 24 biology teachers, 80 examination candidates of 1982, and 4 officers of the Examinations Council. She also analysed items from continuous assessment tests and exercises and projects set by 4 schools. Although the sample used was small, like Meena, she reported that the items were testing facts more than other skills. A high correlation coefficient, (.98) was found between the abilities being assessed in CA and those assessed in the final examination.

Biology is the only compulsory science subject offered and taken by every candidate at O-level in Tanzania. Clearly with the aims of ESR in mind, the content of the biology curriculum, with topics such as, Soil and Soil conservation, Nutrition, Parasites and diseases of man, and Genetics, can make an important contribution to the general education of every candidate and includes issues of concern to all Tanzanians especially in rural life.

Therefore, given the situation (lack of research) that has prevailed in the O-level Biology curriculum and examinations in Tanzania for the last decade, and considering the contributions that biological knowledge can make towards the improvement of life and economy, the present study was carried out.

The study proceeds by making a general brief overview of curriculum and curriculum evaluation in Chapter Two.

Chapter Three makes a general overview of the uses of examinations as a tool for curriculum evaluation.

Chapter Four, onwards, details the case study of the Tanzanian O-level Biology Curriculum and Examinations.

CHAPTER TWO

CURRICULUM AND CURRICULUM EVALUATION

2.10 Introduction

As pointed out in Chapter One, in the 1960's there were a number of movements advocating the need for curriculum change. These movements led to world wide curriculum reforms, notably in mathematics and the sciences, where most of the changes which were introduced were geared towards the development of modern teaching and evaluation techniques, and a re-consideration of what should constitute the curriculum content.

What should constitute the substance of the curriculum on one hand, and how curriculum change should be managed on the other hand, has been discussed by many curriculum theorists, for example, Tyler (1949); Taba (1962); Kerr (1968); Wheeler (1967); Stenhouse (1975); Lawton (1978); and Kelly (1982). It is not within the scope of this study to outline theories of curriculum and of curriculum change. However, the principles of curriculum evaluation, are considered in some detail because the main study presented later focuses on examinations as a tool for curriculum evaluation.

That curriculum evaluation is viewed as part of the curriculum process is evidenced by the inclusion, by curriculum theorists, of evaluation as one of the steps involved in the curriculum process. For example, Tyler (1949) suggested a curriculum analysis focusing on four principal questions:

1. What educational purposes should the school seek to attain?
2. What educational experience can be provided that are likely to attain these purposes?
3. How can these educational experiences be effectively organized?
4. How can we determine whether these purposes are being attained?

These four principal questions may be translated into a four step model: objectives - content - organization and method - evaluation. In this four step model, if evaluation shows that specified objectives have not been attained it must mean that the content chosen or the methods of teaching and organization used were inappropriate, or the objectives were ill chosen.

Tyler's approach has been taken as a model for curriculum theory. The model has developed in at least two main ways: First, the meaning of objectives has been refined, and secondly, the four step model has been elaborated into more sophisticated versions. The model has often been criticised for having one main short coming, that is, it does not take into account "built-in-evaluation", and that its linear approach to curriculum development only allows for summative evaluation (Popham, 1975).

Wheeler has described a curriculum development model which is an elaboration of Tyler's rational curriculum development model. Wheeler's model includes evaluation as a fifth step in a five stages model (Wheeler, 1967 p.30):

1. The selection of aims, goals and objectives.
2. The selection of learning experiences designed and formulated to help in the attainment of the aims, goals and objectives.
3. The selection of content (subject matter) through which certain types of experience may be offered.
4. The organization and integration of learning experiences and content with respect to the teaching-learning process within school and classroom.
5. Evaluation of the effectiveness of all aspects of stages 2, 3 and 4 in attaining the goals detailed in stage 1.

A variant of this model, sometimes referred to as the systems model, has been described by Kerr (1968). Here, evaluation is the fourth step in a model consisting of four interrelated components of:

1. Objectives
2. Knowledge
3. Learning experiences
4. Evaluation.

Kerr and Wheeler strongly agree that the stages (components) are interrelated, interdependent and integrate to form a cyclical process, serving as a feedback mechanism for curriculum modification. This implies that, although the stages may be discussed separately and sequentially from the point of view of their timing and the operations involved, the process should be conceived as a system with inseparable components.

A well known elaboration of Tyler's model is described by Taba (1962) and applied by her as a means of curriculum development. She says, 'curriculum development is a process that requires orderly thinking and order in decision making' (p.12). She goes on to suggest a seven step order for curriculum development:

1. Diagnosis of needs.
2. Formulation of objectives.
3. Selection of content.
4. Organization of content.
5. Selection of learning experiences.
6. Organization of learning experiences.
7. Determination of what to evaluate and the ways and means of doing it.

The above design looks comprehensive and seems to encompass questions of learning methods, instructional techniques and educational evaluation. Although many authors suggest that evaluation is a component of the curriculum process, Taba claims that rarely are plans for evaluation included in the process. She suggests that plans for evaluation should include how the quality of learning can be evaluated to assure that the ends of education are being achieved.

So far, the terms curriculum and evaluation have been used without assigning distinctive meaning to them. In the next section the curriculum in the context of this study is defined. Evaluation is defined in section 2.30.

2.20 Curriculum Defined.

There are several approaches to defining curriculum. The narrowest approach is to use the term 'curriculum' as if it was synonymous with 'syllabus'. In this context it is possible to talk of the 'science curriculum'; or more specifically, the 'chemistry curriculum', or 'physics curriculum' or 'biology curriculum' meaning the specific content brought together and constituting a 'subject'. Defined this way, it means the content of a subject which students are to cover. It is distinguished from the school's organization, teaching methods, and from extra-curricular activities.

Another approach is to define curriculum in a broader sense, extending the subject curriculum. For example, in terms of what educational institutions set out to do, or in terms of what operates in educational institutions with the intention of altering learners' behaviour. In this approach, one can speak of the 'school curriculum' including not only the available subject curricula, but also all the other educational encounters relating to personal relationship, moral attitudes, and social habits offered to students by the school, that is, to include the 'hidden curriculum'.

In this broad sense, the curriculum is wide, and includes all the chosen activities which the school seeks to achieve. According to Oliver it is 'what happens in a school as a result of what teachers do, includes all of the experience of children for which the school accepts responsibility' (Oliver, 1965). Thus curriculum is 'all the learning which is planned and guided by the school, whether it is carried on in groups or individually, inside or outside the school'. (Kerr 1968, p.16).

Other approaches to defining curriculum have been expressed by different curriculum theorists. For example, Gagné takes a psychologist's stand and defines curriculum as 'a sequence of content units arranged in such a way that the learning of each unit may be accomplished as a single act, provided the capabilities described by specified prior units (in the sequence) have already been mastered by the learner' (Gagné 1967 p.23). Gagné's definition implies that a curriculum may be of any length, that is, it may contain any number of units. It is the structure and sequence of the units which give coherence to the curriculum rather than the length of its units. Also, according to the definition, a curriculum is specified when the terminal objectives are stated and the sequence of prerequisite capabilities assumed to be possessed by the students are identified.

Several other curriculum definitions can be found in Stenhouse (1975), Owen (1973), Popham (1975). All the definitions given above have two elements in common: the learner, and the programme.

The learner is the focus in that he is the point of output of the activities set out in the programme. The following definition takes into account the narrow and the broad definition, 'a curriculum is an educational programme. It can be informally organised: what a craftsman teaches an apprentice; or formally organized: what is taught in an instructional film. A curriculum defined in this way, could be a mere lesson, or it could be the curricular programme of a comprehensive high school, or the entire educational programme of a nation. A curriculum may be specified in terms of what the teacher will do, in terms of what the students will be exposed to, or in terms of student achievement' (Stake, 1967a p.4).

However, for the curriculum developer, the broad definition is important for three reasons (Becher & Maclure, 1978). First, the terms of reference have to be wide because the aim is to improve the education which the students actually receive. If he (the curriculum developer) confines himself to the design of syllabuses, he will be neglecting a range of curricular influences.

Secondly, the work of curriculum development is not completed till it has actually reached the classroom and influenced what actually goes on between individual teachers and learners.

Thirdly, curriculum development has to be regarded as the responsibility of a much wider group of people than just the curriculum developers. Thus, the development process ought to be seen in a much broader pedagogic setting.

2.30 Approaches Towards Defining Curriculum Evaluation.

There are several possible approaches towards defining curriculum evaluation. This section involves a brief overview of the development of evaluation concepts and approaches. Prior to the 1960's three elements seem to have dominated evaluative approaches: the use of expert judgement; measurement and Tyler's emphasis on objectives, (Maling-Keeps, 1978).

The use of expert judgement; and measurement are not given detailed coverage in this study because, expert judgement was used mainly to give accredited reports, focusing on criteria such as number and quality of books in the school; number and staff qualification; buildings, for example, libraries, classrooms, laboratories and the standard of their equipment; while measurement was more closely made in relation to IQ and Aptitude tests, which are also not within the scope of this study.

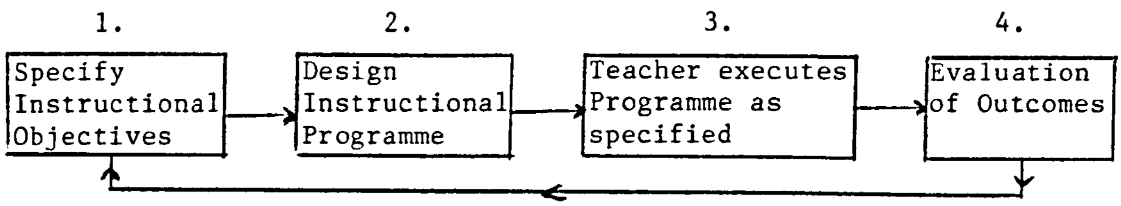
Tyler's Objectives approach.

As far back as the 1930s, Tyler developed his ideas on the need to formulate objectives, define these in terms of student behaviour and then, through measurement, assess the extent to which the objectives had been achieved. (Tyler, 1931, 1949).
With the development of taxonomies, designed to classify educational objectives in terms of observable student behaviour, for example, Bloom (1956)and Krathwohl et al (1964), the systematic analysis of desired outcomes became a real possibility. The devised classification systems covered both cognitive and affective domains, and, although criticized they still remain the most frequently used forms of analysis of educational outcomes.

During the 1940s and 1950s, therefore, it became clear that Tyler regarded testing as the key aspect of evaluation. Tyler's approach is often referred to as the 'Goal attainment model' (Popham, 1975; Jenkins, 1976). The model involves a careful formulation of educational goals, and then, the transformation of these goals into instructional objectives. At the end of an instructional programme based on the formulated instructional objectives, assessment of students is undertaken in order to determine the degree to which the objectives were achieved.

Diagrammatically, Tyler's goal attainment model can be presented as shown in Fig. 2.01.

Fig. 2.01 Diagrammatic presentation of Tyler's Goal-Attainment Model for Curriculum Evaluation.



The goal attainment model has been criticized for its linearity (Popham, 1975), and that it is more suited for the evaluation of, 'motor skills, office procedures and the like' (Jenkins, 1976). However, the model seems, so far, to have provided a useful tool for curriculum evaluation as Jenkins comments:

> 'It would scarcely be an exaggeration to claim that all other schemes endorse, build up, or react to the Tyler model.' (Jenkins, 1976, p.26).

Cronbach: ("Utility" approach).

In the late 1950s, there was a shift of emphasis away from objectivity and validity towards utility. This concept of utility gained increased recognition during the 1960s and 1970s. Amongst the earliest to doubt the completeness of Tyler's objectives approach was Cronbach. Cronbach, (1963), pointed out three major types of decisions for which evaluation data may be used. These are for;

1. Course improvement by deciding what instructional materials and methods are satisfactory and where change is needed through the determination of strengths and weakness which can be maintained, changed, or eliminated.
2. Making decisions about individuals by identifying the needs of the pupils for the sake of planning his instruction, judging merit for the purpose of selection and grouping, and acquainting the pupils with their own progress and deficiencies.
3. Administrative regulations in judging how good the school system is, how good individual teachers are, and any such school factors.

In his view, evaluation is regarded as 'the collection and use of information to make decisions about an educational programme.' (Cronbach, 1963)

Scriven: (The goal-free approach)

According to Scriven, Cronbach's definition of evaluation leaves unanswered the question as to whether evaluation has taken place before judgemental criteria are applied to the information collected. Scriven (1967), argued that the determination of worth or value not only is a central part of evaluation, but such judgement should also be made by the evaluator. He forwarded three notions: evaluation of goals; goal-free evaluation; and formative and summative evaluation. One of his major concerns about the objectives approach was that it left the objectives unexamined. He argued that there was no use in assessing the degree to which a course achieves its goals, if these goals were inappropriate.

In addition to assessing the appropriateness of the goals, Scriven developed the notion of goal-free evaluation where, in direct contrast to Tyler's approach, the evaluator begins by deliberately avoiding the teacher's or curriculum developer's goals. By this approach, observations are not directed towards intended and expected outcomes, but concentrate on actual outcomes. By the end of an evaluation carried out through this approach, important discrepancies between intention and reality may be revealed, and the original goals can be scrutinized in the light of the actual processes and outcomes/products.

Scriven pointed out that evaluation can either be formative or summative. Formative evaluation refers to assessment of worth focused on programmes that are still capable of being modified. The formative evaluator gathers information and judges the merits of aspects of an instructional sequence in order to make the sequence better. Formative evaluation is investigatory, no matter which elements of the curriculum are being examined. Scriven clearly points out that the main purpose of formative evaluation is to serve as a tool for providing feedback to decision makers, planners, teachers, students and as a tool to predict summative evaluation. He then points out that summative evaluation refers to assessment of merit focused on completed plans, purposes and instructional programmes. Thus, summative evaluation should report relative levels of success and failure according to specified criteria, (standards or values), and is not necessarily intended to provide information for subsequent modification and development.

Stufflebeam: (The Context-Input-Process-Product model).

Stufflebeam, in the late 1960s and early 1970s, strengthened the relationship made by Cronbach between evaluation and decision making. Stufflebeam, (1967 and 1971), spelt out the need for evaluations to help in planning, programming, implementing and recycling decisions. His concern for these led him to propose four types of evaluation: Context evaluation; Input evaluation; Process evaluation; and Product evaluation (Stufflebeam, 1971, p.218), each particularly suited for a category of decision type as shown below:

1. Planning decisions would be best served by 'context' evaluation which, by analysing the situation and attempting to relate actual

and desired conditions, would help to provide a rationale for identifying objectives.

2. Programming decisions would best be served by 'input' evaluation on the availability and use of resources, and on such matters as the design of the programme.
3. Implementation decisions which require the kind of information that indicates how things work, and what might go wrong, would best be served by 'process' evaluation.
4. Finally, 'product' evaluation would help in making decisions related to the worth or success of, for example, a new programme, and to considerations of whether to recycle, modify or reject.

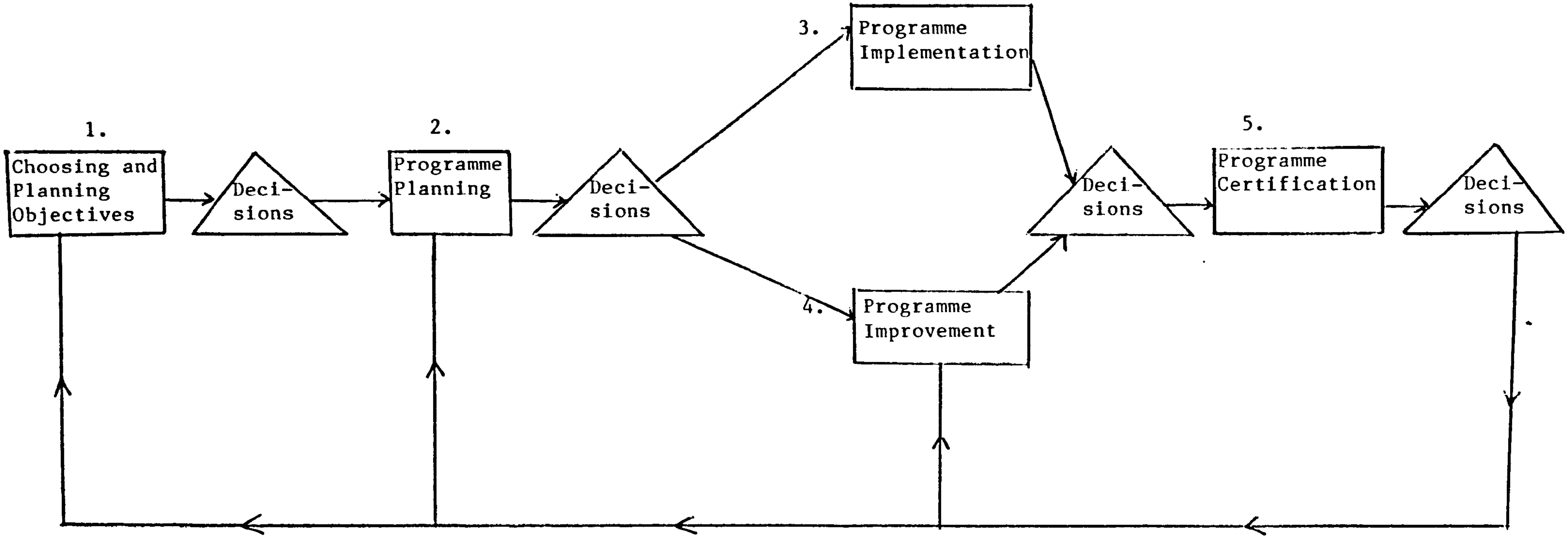
The product evaluation in the Context-Input-Process-Product (CIPP) model, closely resembles Tyler's goal-attainment model in that it involves operationalizing the objectives, measuring criteria associated with the objectives, and comparing the measurements with predetermined standards. As in Tyler's model, product evaluation reports on the degree to which objectives have been achieved. The main short-coming of the CIPP model is that it may 'overvalue efficiency' because it has been perceived as being politically orientated (Jenkins, 1976, p.48).

Diagrammatically, the CIPP evaluation model can be presented as shown in Fig. 2.02.

Stake: (The Countenance model).

Stake (1967b) has been responsible for extending the concept of evaluation. His evaluation model, popularly known as 'Stake's Countenance model' carries the following ideas: First, that any legitimate evaluation should find room for both formal and informal evaluation techniques. Secondly, that both descriptive and judgemental data should be involved in any evaluation. In this respect, he suggests three types of data that should be collected: Data covering the conditions existing prior to the teaching-learning

Fig. 2.02 Diagrammatic presentation of the CIPP evaluation model.



of the programme itself, (antecendent data); data covering the encounters and negotiations, (transaction data); and data covering the outcomes, intended and unintended, (outcome data).

The third idea concerns analysis and portrayal of a wide variety of elements associated with the programme or events being examined. For example, an analysis, according to Stake, should produce a comprehensive portrayal, reporting the ways different people see and feel about the curriculum.

From the previous sections, therefore, the evaluation can be of specific goals, general goals or goal-free; external or internal; of process or product; of an individual or group; of the whole institution or of parts of the institution; formative or summative. In whichever form, and/for whatever purposes the evaluation has been designed, there are two things which seem obvious, that evaluation:

1. will involve gathering information/data.
2. Using the information/data to make decisions.

Thus educational evaluation can be defined as 'the process of delineating, obtaining and providing useful information for judging decision alternatives' (Stufflebeam et.al. 1971, p.43).

2.40 The Process of Curriculum Evaluation.

Through the curriculum process, the elements of curriculum are spelt out. Regardless of the steps involved in the curriculum process, and the evaluation approach used, there are four elements of the curriculum to which evaluation might address itself: The objectives; the content; the method; and the evaluation of student achievement of stated objectives.

Since the curriculum is essentially a plan for helping students to learn, ultimately the central features of curriculum evaluation should address the criterion of effectiveness of learning and attainment of objectives. However, evaluation might be used not just for the terminal assessment of pupil changes, but also to produce independent evidence about the effectiveness of each stage of the course as it is developed (Kerr, 1968). Thus, curriculum evaluation might begin with the concern about the objectives, proceed to content and process, and end with an assessment of their attainment. This approach has been taken in the case study presented from Chapter 4. The assessment system should be considered as 'an essential part of the curriculum and not something divorced from it' (Mathews & Leece, 1976).

2.50 Evaluating a Centralized Curriculum.

2.51 Characteristics of a Centralized Curriculum System.

Tanzania has a centralized system of education. The characteristic features of a centralized education system which are observable in Tanzania, and affect curriculum are:

1. Curriculum development is seen as an integral part of a larger policy of educational reform (ESR). As in most centralized systems, the social and political objectives of educational reform in Tanzania, are determined by the Parliament, and they, (the objectives), in their turn, are part of a larger policy for the transformation of society as a whole. In the case of Tanzania, as was pointed out in Chapter One, the larger policy is the policy for Socialism and Self-Reliance spelt out in the Arusha Declaration (TANU, 1967).

2. In a centralized system, curriculum guidance takes the form of the production of very full documents, for example, syllabuses, teachers' guides, textbooks, which are sent to individual schools and which constitute the public curriculum.

3. Through the published curriculum documents, changes in the curriculum are formally registered and transmitted to the schools. This is achieved easily because curriculum development is one of the functions of, usually, a government department responsible for curriculum administration. As pointed out in Chapter One, in the case of Tanzania, the responsible institution is the Institute of Education.

4. Often, with a centralized system, there are research bodies which are integral parts of the curriculum development which can be called upon when research is required in certain aspects of the curriculum.

There are disadvantages of having a centralized system of curriculum control (Becher and Maclure, 1978). There is always a gap between the public curriculum, (the ideal), and what teachers and pupils actually pursue, (the real); it is not easy to closely supervise what is happening in the schools; there is a danger of diluting the educational content in an attempt to deal with the needs of the changing society; rarely are draft publications tried out and then revised in the light of systematically collected classroom experience, in the sense that it is not easy for individual educationalists to attempt and undertake curricular research and development.

2.52 Problems of Evaluating a Centralized Curriculum.

There are several problems which may be encountered in the course of evaluating a centralized curriculum. First, in a centralized system the goals of curriculum reform are too widely drawn for the evaluator to attempt to measure a manageable number of specific outcomes, or assess a clear-cut set of desired skills, and to persuade himself and others that by doing so he will have provided a reasonably comprehensive examination of the curriculum.

Secondly, following the first problem, the evaluator has to rely on identifying what seem to be the most appropriate and sensitive indicators of whether or not the curriculum is working well, and then on arguing his case indirectly from pieces of partial evidence.

Thirdly, because a centralized curriculum usually has substantial repercussions on a whole range of educational activities, both its supporters and opponents are more concerned to be proved right than to be given a fair picture of what is going on. The issue tends to become heavily politicised, meaning that unless the evaluator is prepared to sell his services as an advocate for one side or the other, his activities may embarrass his sponsors and be seen as a positive threat. In such a system, the importance and weight attached to assessment becomes out of hand. Performance on public examinations is one of the best indicators of the success or failure of a public curriculum and hence analysis of factors which contribute to such success or failure is potentially a powerful evaluative tool.

The third point above is not always specific to centralized curricula but to curriculum per se because it is the case that questions of evaluation present some of the most difficult problems in curriculum studies, in the sense that, evaluation is not simply concerned with how well students have learnt something, but also with questions of justification and unintended consequences of learning.

The problems of evaluation, according to House (1973), lie in the origin of evaluation. He says:

> 'Evaluation is always derived from biased origins. When someone wants to defend something or to attack something, he often evaluates it'. (House, 1973, p.3.)

2.60 Evaluation of School Science Curriculum.

Although little has been done with regard to evaluation of school science curricula in Tanzania, extensive work, covering the use of a range of evaluation models, has been done in the evaluation of school science curriculum in other parts of the world. Of main interest in this study are those done in school biology curricula, such as: evaluation of overall student achievement (Kelly, 1971); evaluation of materials (Monger and Kelly, 1974); evaluation of specific objectives (Amos 1967; Duckworth and Hoste; 1976, Kelly, 1972); evaluation of learning difficulties and conceptual demands (Shayer, 1974; Okeke and Wood-Robinson, 1980; Johnstone and Mahmoud, 1980; Simpson and Arnold, 1982); study of examination performances (Willmott and Hall, 1975).

These studies and others are reviewed in more detail in later chapters. In the next chapter, the role that examinations can play as a tool for curriculum evaluation is discussed in more detail. Later, evaluation of the Tanzanian O-level Biology Curriculum using examinations performance as one of the tools, is the focus of the case study presented from Chapter Four onwards. Several of the ideas from the evaluation approaches described in the earlier sections of the present chapter have been utilized. For example, Scriven's argument that there is no use in assessing the degree to which a course achieves its goals without examining the goals themselves has been taken into consideration and the curriculum objectives have been evaluated in terms of their importance as O-level course objectives and examination objectives.

Tyler's goal-attainment or the 'product' evaluation of the CIPP model has been used to evaluate the attainment of the curriculum objectives as measured by the performance on the examinations.

Process evaluation of the CIPP model has been used to gather information on matters relating to the conduct of practical examinations and the process of continuous assessment in general.

In addition to evaluating the objectives and attainment of the objectives, the content of the curriculum has been studied in an attempt to identify areas both of perceived and actual difficulties using both 'soft' and 'hard' data gathered through both indirect and direct approaches.

The details of the evaluation procedures and their outcomes are presented separately in the relevant chapters.

CHAPTER THREE

PUBLIC EXAMINATIONS, A TOOL FOR CURRICULUM EVALUATION

3.10 Examination defined.

Harris defines examinations in educational usage as 'a series of questions or tasks designed to measure the knowledge or skills of an individual' (Harris, 1960, p.1502). Examinations can also be defined from their organizational aspect. Defined this way, "an examination is a method of assessing candidates on the basis either of special tests or on study of the school record or a class teacher's assessment, or of a combination of these procedures" (Agazzi, 1967, p.66). Basically, the above definitions indicate that examinations are bound together by a single purpose, that of measurement.

3.20 Why Examine?

The data obtained from examinations can serve a variety of purposes. Hartog (1938) talking about examinations in England pointed out that "the greatest weakness of the 'present' system of examinations in England, and in many other countries, is the failure to define with any degree of exactness the purpose of each examination and each examination paper and other test forming part of an examination" (p.5). He then points out seven possible uses of examinations: to arrange candidates in order of proficiency; to test progress towards the attainment of utilisable skills; to test the efficiency of schools; to give a sanction to some course of teaching; to test intelligence; for vocational guidance; and to test culture.

Morris (1961) classifies the purposes of examinations under four headings: as a device for maintaining standards; as a device for stimulating effort; as an administrative device; and as a tool of social engineering. (Morris 1961, pp.1-26).

Agazzi (1967, p.74) points out that the alleged functions and real or supposed uses of examinations are numerous. In a summary he presents at least five of those purposes. These are to provide teachers with information about their pupils, thus enabling them to give more and effective help; pupils with an indication of their progress; data for guidance; parents and the public with reliable information on the overall results of the educational system; a sound basis on which to adjust curricula to meet the needs of a constantly changing society; and to provide data for selection.

Pilliner (1968) lists somewhat similar purposes to those of Morris, in terms of the student; the teacher; and the society at large. 'To the student they are a stimulus and a goal; to the teacher they feed back information about the effectiveness of his teaching and hence serve as feedback mechanism; to the society at large they furnish a guarantee of competence in those examined to perform the tasks demanded of them by the jobs or professions they take up' (Pilliner, 1968 p.167)

Kelly (1971) in a reappraisal of examinations points out that the major roles of present day examinations are as measures of achievement which provide criteria for judging standards of performance by students (and hence teachers and schools) and as targets acting as incentives for their efforts; as a selection mechanism; and as a means of evaluating the instrumental social and economic roles of the educational system. (Kelly, 1971 p.119)

Smith and Adams (1972) categorize the use of examination results under four categories: for the teacher; the school administrator; the guidance counsellor; and the student. In all four cases, they point out that the examination results are a 'feedback mechanism' and help to direct decision-making.

Lewis (1974) classifies the purposes of examinations into two groups. Those served by school examinations, and those served by external examinations. He says that school examinations are devised more with the intention of promoting uniformity in the classroom than with assessing individual differences. They have several purposes: "for example, to ensure that certain standards have been reached, to provide incentives to pupils, and 'feedback' information to teachers, and to allocate pupils to courses of instruction ('stream' or 'sets') from which they are likely to benefit" (Lewis, 1974. p.91).

Externally organized examinations - such as those for the General Certificate of Education (GCE) and the Certificate of Secondary Education (CSE) in Britain have several social functions. He says such examinations help to decide whether competence in a given field has been attained, and act as selection measures for higher education.

Ward (1980) lists similar purposes to those of Pilliner: for the student, for the lecturer or college, and she categorizes the society at large into two sub-groups; the employer, and the public. To the employer, examinations are a measure of employability; and to the public they are a measure of safe guarding public safety and health, and a means of ensuring adequate standards.

Thus one can agree with Agazzi (1967) that the purposes of examinations are numerous. But looking through all the works cited, the purposes of examinations can be summarized by saying that we examine to: motivate, diagnose, select, report, compare, predict, and to evaluate. The final element in the examination process is to make decisions. Tests are useful only if they can help in making decisions. Each purpose of a test discussed above can be traced to a need to make a decision.

An examination may possibly be designed to serve one or more of the above functions. It is very important to be as specific as possible about the essential functions that a particular examination is to serve. For it is the use that will be made of the examination outcomes which will determine the type of the examination, and the kind of items to be used. Unless the purposes of an examination are defined clearly, it will be difficult to say whether a given question in it is relevant or irrelevant.

3.30 Uses of Examinations to Evaluate Curriculum.

Much has been reported on modern curriculum evaluation, but there seems to have been suprisingly little use made of examinations as a means of evaluating curriculum in biology or in any other curriculum area. There are three basic conditions to be met in order that examinations can be used to evaluate curriculum (Mathews and Leece, 1976).

1. The specification of the purpose; the content; and the processes of the curriculum must be made explicitly and quantitatively in the curriculum itself.
2. The assessment system, should match the purpose, content and processes of the curriculum both qualitatively and quantitatively.
3. A substantial part of the curriculum must be studied and assessed.

Extensive use of examination results was made in the formative evaluation of the Nuffield A-Level Biology Project. Different types of examinations were used in evaluation studies carried out on this project (Lister, 1969; Kelly & Lister, 1969; Kelly & Dowdeswell, 1970; Kelly, 1970, 1971).

Tamir & Glassman (1970; 1971; 1972) designed practical examinations and used them to evaluate the BSCS biology project in Israel. The details of their work are discussed in Chapter 8 when practical assessment is considered.

Johnstone and Mahmoud (1980, 1981) conducted an evaluation of the school biology curriculum in Scotland. They used, as part of their data source, examination reports which pointed out strengths and weaknesses exhibited by candidates. A comparison was then made between the areas which were identified as difficult and topics which pupils and teachers perceived as being difficult.

Two studies are particularly relevant to the present work, they are those of Mathews and Leece (1976) on the evaluation of Nuffield Chemistry and Kelly (1971a): on the evaluation of the Nuffield A-level Biological Science Project. The latter, entitled 'Overall Achievement of Students', is reported in summary below, because it points out issues that should be considered when using examinations to evaluate curriculum.

Three considerations guided the strategy of the project with respect to examinations:-

1. The examinations were developed both to evaluate the effectiveness of the course and to measure attainment in a way suitable for selecting students for higher education.
2. The examinations were intended to reflect the work of the course not only in terms of the subject matter covered and the abilities the students used, but also through the modes of thinking employed and the involvement of students in activities similar to those undertaken in the course.
3. It was contended that by encouraging teacher participation in the examination process, greater validity could be obtained; that is, the objectives and methods of the course would be better understood and absorbed into teaching; a more immediate and sensitive feedback concerned with the relation between examinations and the course would be possible; and teachers could more quickly acquire the techniques of examining and incorporate them into their own system of class testing and diagnosis of student performance.
(Kelly, 1971b).

For evaluation purposes, the examinations assessed the students' knowledge of a range of subject matter and their facility in using a number of abilities assumed to be necessary for understanding biological science within the context of the type of activities undertaken during the course.

The examination was divided into six sections carrying three forms of examination papers with different kinds of items. Sections A, B, C and D were written papers with multiple-choice; matching, continuous prose; and short-answer items. Section E was a project which was assessed by the teachers and moderated through the use of final reports. Section F was Practical course work which was assessed by the teachers.

It is important to note that the design of the examination was based on the activities associated with the wide range of educational objectives of the course.

Teacher participation is an important element in the achievement of curriculum objectives and in curriculum evaluation. In the above project, the vast majority of those concerned with writing items, marking, assessing and moderating the examinations were practising teachers.

Kelly points out three things which were gained by the participation of teachers; an understanding of the objectives and techniques of the examination was acquired by the teachers; a bank of examination items was produced for mutual use by teachers, and an intimate feedback system was established, providing information to help the classroom practice of the teacher and acting as a continuous source of curriculum development.

The examinations acted as an evaluation instrument in that they provided a profile of the students' performance in the areas examined. The performance was then used to judge the efficiency of the methods and materials and hence provided information on which judgement could be made and potential adjustment effected.

3.40 Use made of examinations and examination procedures in Tanzania.

The use made of classroom test results, school examinations and National examination performance in Tanzania is summarized below. These are considered under three headings: the student, the teacher/school; and the public.

For the student:

Classroom test results and examination performance are used as:-

a. an incentive to study during the course.

b. a means of indicating areas on which a student needs further study during a course and especially towards the end of the course before the final examinations are taken.

c. a recognition of his knowledge or skill and of the effort he has put into the course.

d. a means of entry to a profession. For example, in order to be selected for entry into a nursing career, a candidate must obtain, among other passes, a credit pass in biology.

e. a means of entry to a higher course or a later stage of the same course. For example, in order that a candidate be enrolled in A-level courses, where biology is offered as one of the subjects, the candidate must obtain a credit pass in biology and in the other subjects offered in that particular combination.

Purpose (d) and (e) above are based on the results of the National O-level Examination results which 'serve a dual purpose, they are both terminal and promotional.' (Education, Ministry of, 1980, p.17).

For the Teacher/School.

Tests and examinations in Tanzania are part of the teaching process. They are used by the teacher/school as:

a. a tool to indicate (diagnose) areas on which further tuition or practice is needed during the course.

b. a tool for personal evaluation of the success of the teaching, in relation to the aims of the course and in comparison with other teachers and other schools.

c. an evaluation of the success of the course and a source of information relevant to further curriculum development. Such information is passed on to the Institute of Education, which is the National Curriculum Centre, for consideration through panel meetings, seminars, or research projects.

d. a guide to the preparation of student progress reports sent to each parent, for information and comments, at the end of every year.

e. a means of fulfilling the requirements of the National Examinations Council which demands a continuous assessment mark for each candidate for each subject to be supplied by the teacher/school. The continuous assessment mark forms part of the final assessment for the award of a Certificate of Secondary Education.

For the Public.

The performance on the Certificate of Secondary Education Examination ('O' - level) are used by the public as:

a. a measure of employability. For example, certain employers, that is The Ministries, Parastatal organization, and private companies, will only consider for employment, those who have passed and obtained credit passes in certain subject combinations only. For example, a pass in the science subjects, especially with the combination of biology, physics, chemistry and mathematics, will almost guarantee a candidate preferential employment in the Ministry of Agriculture, Health, Education and into the industrial sector.

b. a measure to ensure adequate standards of work.

c. an evaluation of the quality of education being financed by the tax payer.

To fulfill the purposes outlined above, a combination of different types of examinations and examination procedures could be employed. Types of examinations which can be used to evaluate curriculum are discussed below.

3.50 Types of Examinations.

One of the purposes of examinations listed in the previous pages is to evaluate the curriculum. The need to evaluate curriculum has been discussed in Chapter Two of this study. The discussion that follows centres around the types of examinations which can be used in curriculum evaluation and the kind of information they yield.

Examinations can be classified into three main categories according to assessment procedures; types of tasks demanded of candidates; and types of questions.

Category I: based on the assessment procedure.

Examinations may be internally or externally organized. Internally organized examinations are compiled and conducted by the students' own school and teachers. In the British CSE and O-level examinations, such internally organized examinations are referred to as Mode III. The examinations are based on a syllabus drawn up by the school and validated by the examining board.

Externally organized examinations are compiled and conducted by an outside examination body. In England two kinds of external examination procedures exist: Mode I where the examination is compiled and conducted by an examining board based on the boards' examining syllabus; and Mode II where the examination is compiled and conducted by an examining board but based on the school syllabus.

In Tanzania, the school syllabus is the same as the National Examinations Council's syllabus for each examined subject. Both internal and external assessment procedures are practised. Internal assessment is mainly carried out through continuous assessment (CA). For details on CA see Chapter 9.

External assessment is carried out by the National Examinations Council of Tanzania - the sole examination body which has been vested with powers to examine.

Category II: based on form of task.

Written, practical and oral examinations are the best known forms in this category. Of these, the written examinations are most popular. Oral and practical examinations are generally included in subjects where such skills need be assessed, for example, oral examinations are appropriate in languages, while practical examinations are appropriate in the sciences.

In recent years, variations on these forms of examinations have begun to be tried out. One of the most interesting variants is continuous assessment of course work. This is an assessment of work performed by the candidates during their school course years. This may include all three types of examinations mentioned above - written, oral and practical.

Practical examinations constitute separate papers in the O-level science examinations. Biology practical examinations of Tanzania constitute the part of the research reported in Chapter 8.

Category III: based on type of question/items.

The classification of examinations according to the type of question has been discussed in detail by many specialists in educational measurement. The details of item characteristics, methods of construction, advantages and other related issues are extensively treated in works, such as those by Macintosh & Morrison (1969), Thorndike and Hagen (1969), Cronbach (1970), Gronlund (1971), Nunnally (1972), Ebel (1972), Thyne (1974), Macintosh & Hale (1976). In this study, a brief summary of item types and the demands made by them on students is given below.

Basically items can be classified into two major forms: objective items and subjective items. Macintosh & Morrison (1969) define objective items as 'those items which are asked in such a way that for each of them there is only one predetermined correct answer' (p.9). It follows, therefore, that subjectivity in the marking, or scoring of the test is eliminated or is minimal and that it can be marked even by an individual without any knowledge of the subject. The objective items include: Matching; fill-in (completion); true-false; and multiple-choice. The latter is probably the most commonly used type of objective item.

In subjective items, 'what is to be accepted as the correct answer is a matter of opinion' (Thyne, 1974, p.179). Subjective items include: short-answer questions; extended-answer questions; and essay questions. Table 3.00 is a summary describing the nature of the different item types and the responses to them expected from the candidates. All these different types of items are used in Tanzania for assessing student achievement in O-level biology. Performance on the different types of items by students generally,

Table 3.00 Types of items and required candidate response.

Type of Item	What is presented to the candidate	What is required from the candidate
True-false	statements	to mark either true or false
Fill-in	incomplete statements or questions	to supply one term, name, date or phrase which either completes the statement or answers the question
Matching	presented with two lists of names, facts, or principles	to match corresponding options
Multiple-choice	facts, principles, questions, problems, and alternative options	to choose one alternative response.
Essay	problems or questions	to plan his own answer and to express it in his own words
Extended or short answer	problems or questions	to supply the solution or answer in his own words

by different abilities and by sexes, has been reported in a number of research works, for example Dunn and Goldstein (1959); Millman (1966); Fairbrother (1975); Mathews and Leece (1976); Murphy (1978); Harding (1981).

The rest of this study centres on curriculum evaluation, using examination performance as the main source of data. The curriculum which has been evaluated is the Tanzanian O-level Biology course. The study includes an examination of the performance on the different types of items included in the different components of both the Theory and Practical examinations (See Chapter 5).

In order that examinations serve the purpose they are meant to serve, they must show a certain degree of reliability and validity. Reliability and validity of examinations are discussed in more detail in Chapter 6, and Chapter 7 respectively.

3.60 <u>Procedures for setting and marking the examinations in Tanzania.</u>

Teachers' involvement in public examinations in Tanzania is of two types:

1. Work is undertaken by teachers acting as setters, moderators, invigilators, markers or acting as members of the Examinations Council's panels committees and/or Curriculum panels committees.
2. Teachers undertake the assessment of their own pupils in both the Continuous Assessment (CA) scheme, and in the final examination by acting as markers.

Work undertaken in both categories above is important and essential for the Examinations Council to be able to function. In most cases, and as far as the individual teacher is concerned, this work cannot wholly be done voluntarily especially the Continuous Assessment (CA) element. Teachers preparing pupils for the examinations often have little option but to carry out the CA activities which involve setting, marking, and recording weekly exercises, monthly tests, terminal examinations, and yearly projects laid down by the Examinations Council as part of the final award requirement.

Both categories of involvement may involve teachers in absences from school, at weekends, and in the holidays. Staff absences on examinations duties cause difficulties in the schools. But since it is necessary to have the examinations, it has been considered that teachers should play an active role in running the examinations. Therefore absences are tolerated by the parties concerned.

Another problem arises from the fact that, to some extent, teachers have to volunteer for the involvement in the marking of the final examinations. Usually the marking of these examinations takes place during the long school holidays.' All markers meet at marking centres where the marking begins by a detailed discussion of the proposed marking schemes. It is only after reaching an agreement on the marking schemes that individual scripts are dealt with. Occa-sionally there is lack of co-operation on the part of some individual teachers who even though their names would have been proposed by their heads of school, and who would also have indicated their willing- ness to participate, fail to turn up at the marking stations making it too late for replacement to occur.

One of the more general problems in using examinations to evaluate the curriculum is related to coverage. Not all of the goals of education can be measured effectively in an examination situation. Any examination constitutes only a sample of items from a much larger potential population of items, and therefore provides a very incomplete definition of the goals of achievement. For this reason, an examination may be regarded as a sample of items which constitute one operational definition of the testable goals of instruction in the area under consideration.

Another general problem is related to what people think and believe public examinations to be. Examinations are 'used as weapons in a struggle for dominance between groups with conflicting interests', i.e. " 'the discipline' or profession, parents, employer (society) over another group (the students)" (Rowntree, 1977, p.3).

CHAPTER FOUR

A CASE STUDY OF THE 1983 'O-LEVEL' BIOLOGY EXAMINATIONS IN TANZANIA.

4.10 Introduction.

The educational context from which the present research springs has been described in Chapter One. The research is considered to be timely both in terms of the historical development of the curriculum in Tanzania's centralised educational system and, particularly, because of the paucity in the volume of research which has been conducted in science education in Tanzania.

At all stages, the research conducted has been exploratory in that it has followed no pattern which has been established previously in Tanzania, or it is believed, in East Africa generally. One of the aims of the work conducted has been to identify, at all stages, key areas of difficulty or concern which may be useful for future researchers and curriculum developers in Tanzania. However, since the study provides an in-depth analysis of a public examination which is used to provide summative assessment of pupils and on which a range of agencies base selective judgement, the work has applicability in parts of the world where similar procedures are employed.

4.11 Specific Objectives of the study are as follows:-

1. To summarize the curriculum objectives; examination objectives; and topics and sub-topics of the curriculum.
2. To identify areas which are perceived as difficulty by teachers and students, and to compare teachers' and students' perceptions of these areas.
3. To carry out a detailed analysis of the 1983 O-level biology examination with the intention of identifying topics of 'real' difficulty. This will allow matching of perception of difficulty with performance in the examination.
4. To determine the reliability of the 1983 O-level biology examination.

5. To determine the importance accorded to the objectives, by teachers and panel members, as course objectives, and examination objectives, and to compare the views of teachers and panel members with 'actual' weight given to each objective in the examination.

6. To study some aspects of the Practical examination including an indication of the abilities which teachers and panel members feel can best be examined through practical examinations, and to compare the teachers' and panel members' views with 'actual' examination practice.

7. To identify the content of the examination questions in terms of the curriculum objectives, and topic areas being examined, and to investigate the content validity of the examinations.

8. To examine the construct validity of the examinations by analysing the performance on each examination question and part question in order to determine the precise nature of the topic areas and abilities tested and the extent to which a 'practical mode of performance' seems to occur in these examinations. That is, investigating whether 'theoretical' and 'practical' biology examinations represent separate modes of performance.

9. To make comparisons of girls' and boys' scores on the Theory, Practical, and overall performance, in order to investigate whether the sex of candidates, often assumed to be a 'construct', is reflected in biology examination scores to the extent of producing differences in performance in the Tanzanian examinations.

10. To study some aspects of Continuous Assessment (CA) in Tanzania and to examine in detail the relationship between CA scores and Final Examination (FE) scores in O-level biology using the 1983 CA and FE scores.

11. To make comparisons of girls' and boys' scores on CA and FE.

12. To investigate whether 'theoretical mode of performance'; 'practical mode of performance'; and 'CA mode of performance' constitute separate constructs.

4.20 The Research Strategy.

The roles of the Institute of Education, Ministry of Education, and

National Examinations Council in devising, monitoring and examining the secondary school curriculum have been described in Chapter One. The O-level biology curriculum materials and examinations, also mentioned in Chapter One, which constitute the basic documents for the rest of this study are:

1. The O-level Biology Syllabus (1976);
2. The Teachers' Guide to the Syllabus (1978);
3. The Examination Papers for 1983;
4. Examination performance records (1983);
5. Continuous Assessment record (1983).

A short description of the first three documents follows.

4.21 The 1976 Syllabus.

This is a four year syllabus, the contents of which are presented in two columns, a topic, and notes column.

Under the topic column is presented a list of the main curriculum topics to be covered through instruction throughout the course period of four years. Each main topic is further broken down into sub-topics.

Under the notes column appears amplification in the form of suggestions to be stressed during instruction. (See Appendix I).

4.22 The Teachers' Guide (1978).

The teachers' guide, an elaboration of the syllabus, is presented in four columns: topic outline, objectives, suggested pupil activity, and reference.

Under the topic outline column are listed the curriculum topics and sub-topics similar to those in the syllabus. In addition all that is listed under the notes column in the syllabus is listed under this column.

Under the suggested pupil activity column are listed pupil/teacher activities for the achievement of the instructional objectives. It is under this column that some teaching materials, and experiments are suggested.

The reference column carries a list of suggested reference books. (See Appendix II).

4.23 The Examination Papers.

At the end of the four year biology course, each candidate sits two separate examination papers. These are a 3 hour long theory paper, and a 2 hour practical paper. These examination papers are sat on two separate nationally determined days.

The 1983 theory paper consisted of three sections. Section A consisted of three essay type questions. The candidates were required to choose and answer only one of the three questions. All other parts of the examination are compulsory.
Section B consisted of three questions, each with ten items. One of the questions was a multiple choice item with four alternatives. The second was a matching item, and the third a true-false item. Section C had five short-answer questions. Each question with items, and some items with sub-items. (See Appendix IIIA).

The two hour practical paper consisted of three questions aimed at testing experimental and observational investigational skills. The first question required the candidates to do a series of experiments (the "experimental" section). The other two questions were short-structured questions and the candidates were required to identify, draw and/or comment on given specimens (the "observational" section). All questions on the practical paper were compulsory. (See Appendix IIIB).

The systematic development of the biology examination and the incorporation of different forms of assessment procedures have been discussed in Chapters One and Three. The incorporation of different forms of assessment items is not a feature unique to Tanzania. It is seen in public examinations elsewhere, for example in U.K's CSE and GCE and new GCSE proposals. It is believed that such a variety of items enables students to display their knowledge to the best advantage and to encourage teachers to teach the subject confidently, with the knowledge that if she does so adequately, the examination scores will fairly reflect her pupils' relative abilities in the subject.

4.30 <u>The Curriculum Objectives and the Topic Areas</u>.

<u>Identification of Curriculum Objectives</u>.

The syllabus and the teachers' guide do not include a short list of the curriculum objectives and the curriculum topics. The first task in this study was to analyse the syllabus and teachers' guide so as to identify:

1. The curriculum objectives,
2. The curriculum topics to be covered over the four year biology course.

Using the information listed in the column 'objectives' in the teachers' guide, a preliminary list of 33 objectives was drawn up. Through a further analysis, which involved collapsing objectives into fewer categories, a shorter list of 13 objectives, was obtained. For example, in the original list of 33 objectives, objective 10 and 24 were: ability to show knowledge of biological facts; and ability to show knowledge of biological terminology respectively. These were thought to be so similar that they could be combined to produce objective No1 in the list below. The 13 objectives follow:

Ability to:-

1. Show knowledge and understanding of biological facts, concepts, principles and theories.
2. Comprehend biological texts.
3. Apply biological principles to real life experiences.
4. Plan experiments.
5. Follow instructions in carrying out experimental and other biological procedures.

6. Make and record observations accurately.
7. Interpret data to draw conclusions and make inferences.
8. Draw biological diagrams accurately.
9. Analyse and solve problems.
10. Formulate hypotheses from numerical data, graphs, tables and other evidences.
11. Explain and discuss.
12. Write accurate reports of experiments and other biological procedures.
13. Use manual skills to set up and operate laboratory apparatus.

Identification of the Topic Areas.

Two documents formed the basic reference: the syllabus, and the teachers' guide. It was necessary to prepare a summary of the topic areas of the syllabus to include in the questionnaire for pupils and teachers. A comparison of the topic areas as listed in the syllabus and in the teachers' guide was made on the bases of similarities, differences, and sequence of topics between the syllabus and the teachers' guide. It was noted that the topic areas in the syllabus and in the teachers' guide are basically the same. The difference between the two documents is that the teachers' guide elaborates the topics and suggests classroom activities for teachers and pupils.

A difference was also observed in the sequence of topics. The final list given below is based on the sequence in the teachers' guide. This decision was based on the assumption that since the teachers' guide is meant to help teachers, it most likely forms the reference document followed more closely by most teachers. The teachers' guide is also a more recent document. The main topic areas which were identified as constituting the four-year course are:-

1. Classification.
2. Cell structure and function.

3. Soil and plant growth.
4. Nutrition in plants and animals.
5. Respiration in plants and animals.
6. Transport system in plants and animals.
7. Excretion and osmo-regulation in mammals.
8. Control and co-ordination system in organisms.
9. Movement.
10. Growth and development.
11. Reproduction.
12. Genetics.
13. Evolution.
14. Ecology and interdependence of organisms.
15. Parasites and diseases of man.
16. Man and his natural resources.

It is the summaries of the curriculum topics and of the objectives that formed the basis for the construction of the instruments used in this study. Biology teachers' and curriculum developers' opinions on whether these summaries constitute a valid list of the objectives and topics is discussed in Section 4.60 which deals with the pilot study procedures.

4.40 Instruments used in the research.

For curriculum evaluation by examinations, there are three possible data sources: the students; teachers and examiners.

1. Data from written examinations come indirectly from students and directly from examiners.
2. Internal assessment (continuous assessment) data comes directly from teachers and indirectly from pupils.
3. Information from questionnaires or interviews can be direct from students, teachers or other sources such as curriculum planners. Both direct and indirect data were collected in this study.

4.41 Pupil opinion questionnaire (Appendix IVa)

As a preliminary step in a much larger study involving the evaluation of the school biology curriculum in Scotland, Johnstone and Mahmoud (1980), determined the relative difficulty with which different topics were perceived within the biology syllabus of the Scottish Certificate of Education Examination Board (SCEEB - 1974).

Johnstone and Mahmoud applied a 'direct approach method' to gather data from pupils, teachers and university students using an opinion questionnaire with a list of 15 major topics, outlined in the syllabus.

In the present study, a similar approach is adopted. However, instead of presenting the pupils with a list of the major topics only, as Johnstone and Mahmoud did, the major topics were broken down further into sub-topics. It was hoped that the sub-topics would point out more specifically the areas intended to be covered during the course of teaching and learning. In so doing, it was hoped to increase the reliability of the responses and to pinpoint particular areas of difficulty within large topic areas.

The list of topics and sub-topics was presented in an opinion questionnaire administered to the Tanzanian 1983 O-level biology course graduates. The pupils were required to give their opinion about the ease or difficulty of the topics on the bases of the following categories:-

Easy	=	I understood it first time I was taught it.
Average	=	I had to work hard at it, but now I understand it.
Difficult	=	I have worked hard at it, but I still do not understand it.
Not covered/ Not taught	=	I have not covered/I was not taught this topic. (See appendix IVa).

4.42 The biology teacher opinion questionnaire (Appendix IVb)

This questionnaire was designed to identify any problems which may affect the teaching and learning of O-level biology. It has six sections.

Section I sought general information about the school and the fourth form biology teacher who was the respondent to the questionnaire.

Section II is a duplicate of the pupil questionnaire. It differs from that presented to the pupils in that, while the pupils expressed their views of learning difficulties, the teachers were to express their opinions on the degree of ease or difficulty in teaching the topics on the bases of the following categories:-

Easy	= Students understood it first time I taught it.
Average	= I had to work hard at it, but now they understand it.
Difficult	= I have worked hard at it but still they don't understand it.
Not covered/ Not taught	= I have not covered/not taught this topic.

The main purpose for the inclusion of this section in the teachers' questionnaire was to compare teachers' responses with the student responses. That is, a comparison of the teachers' teaching difficulties with pupils' learning difficulties provides more interesting information than either separately about the curriculum and any potential 'mismatches' which might be identified could be important foci for future work.

Section III and Section IV consist of the list of curriculum objectives prepared by summarizing those listed in the teachers' guide.

Amos (1967), in trying to determine the place of scientific method in biology courses leading to the General Certificate of Education examination at ordinary level, in England and Wales developed a list of 19 objectives of school biology. He then circulated his list of objectives among teachers in the North of England asking them to rate the importance of each objective according to the way they taught the subject for the GCE; would teach the subject in an ideal situation; and the way they thought the O-level biology syllabus of the GCE demanded.

In an almost similar study, Duckworth & Hoste (1976) circulated a list of 15 subject areas and 14 curriculum objectives to item writers and scrutineers, and to teachers to get an opinion of the importance of the subject areas and curriculum objectives in teaching and examinations. The subject areas and curriculum objectives were identified by means of a scrutiny of CSE and GCE examination board syllabuses, review of research and discussions with teachers. The above studies, and others, on rating, ranking, and comparison of objectives are further reviewed in Chapter 7.

Therefore, in Section III and Section IV of the teacher questionnaire, teachers were asked to give their opinion by rating and ranking the importance of each objective in teaching and the examinations respectively. The same list of objectives was presented to biology panel members (curriculum developers) to rate and rank-order for comparison with the views of the teachers.

Section V of the teacher questionnaire sought their opinion on the biology practical examinations; including their indication of the abilities which can best be examined through a practical examination. The last section, Section VI, of the teacher questionnaire sought the opinion of the teachers on continuous assessment procedures in Tanzania.

4.43 The biology panel members' questionnaire. (Appendix IVc)

The biology panel members constitute the functional unit of the Institute of Education responsible for the development of the biology curriculum.

Members of the panel comprise biology teachers, subject specialists from the University of Dar-es-Salaam, representatives from the Ministry of Education, Agriculture, and Health. The secretary to the panel is the biology curriculum developer, resident at the Institute of Education.

The main purposes of the biology members' questionnaire were two fold: first, to seek their opinion through rating and ranking of the relative importance of each of the summarized objectives in the instructional and examination processes; and secondly, to seek their views on whether the developed list of objectives constituted an adequate summary of the objectives outlined in the Teachers' Guide (1978); and to suggest additional objectives if the list was adequate.

Section III sought their opinion on the practical examination. Their responses on the importance of the objectives in the teaching-learning situation on one hand, and examinations on the other hand, provided comparative data to those derived from the teachers' responses.

4.44 Examiners' Analysis Form I, II, III. (Appendix IVd)

In the present study, two approaches to classifying the examination questions, using the derived curriculum objectives and topic areas of the Teachers' Guide have been employed.

1. A classification of the questions in the examinations was developed in which the Tanzanian O-level Biology Curriculum objectives and topic areas were used as the criteria without reference to other systems of classification.
2. A classification of the questions was elaborated in which the first classification was related to another system of classification, in this case to Bloom's Taxonomy of Educational Objectives (Bloom, 1956- See Appendix IVe).

The examiners' working materials included three analysis forms, a detailed list of sub-topic areas similar to the one presented in the pupils' opinion questionnaire, a short description of Bloom's Taxonomy of Educational Objectives, and the 1983 examination questions.

The main purpose of this part of the research was to identify the content of the examination questions in terms of the curriculum objectives, and topic areas being examined, which could be used to investigate the content validity of the examination.

Analysis Form I consisted of a list of the objectives prepared by summarizing those listed in the Teachers' Guide (1978).

Analysis Form II consisted of a summary of the topics as outlined in the Syllabus (1976), and elaborated in the Teachers' Guide (1978).

Analysis Form III consisted of a grid matrix on which were listed the 13 objectives on one axis, and Bloom's Categories of Cognitive Educational Objectives on the other axis. (See Appendix IVd).

The examiners were required to study the 1983 biology examination papers; both the theory and the practical papers and then

I. Using the analysis Form I, to classify each examination question according to the objectives which they felt the question or item was testing. Each item or complete question could test either a single objective or several objectives. When they felt that several objectives were being tested, the item or question number was entered in the grid alongside each objective tested.

II. Using analysis Form II, and the elaborated list of topic areas, to classify each examination question or item according to the topic area they felt the question was testing.

III. Using analysis Form III to classify the 13 objectives, as far as possible, according to Bloom's Taxonomy of Educational Objectives. (See Appendix IVe).

The examiners in this study were a selected group of experienced biology examiners from England and Wales. A total of eight examiners were involved. The Tanzanian examiners were not included because they constitute biology teachers and panel members, and are responsible for setting, moderating and marking the examinations. For this reason, an independent external group was preferred.

4.50 <u>The pupil population and schools sampled</u>.

There are two categories of secondary schools in Tanzania, the public, and the private secondary schools. At the time of this study, there were ninety public secondary schools, that is, Government controlled, (the exact opposite meaning of public schools to that in the U.K.) and over seventy private secondary schools. In 1983, the total number of candidates who registered for the Certificate of Secondary Education Examination (O-level) from 85 public schools was approximately 9760. The target population for the study, therefore, were the graduates of the O-level biology course for 1983 from these schools.

4.51 Sampling the Schools.

For the purpose of assessing progress in education, we need to know what all, or almost all, of the children are learning; what the most able are learning; as well as what is being learned by the middle or average ability children (Tyler, 1967, p.14). For this reason, the schools sampled included representative schools from the upper, middle, and lower achievers.

In this study, as in any other research, it was almost impossible to include all elements, or variables operating within the population in the sample. Variables which operate within Tanzanian secondary schools, and which might be related to differences in academic performance are, for example, sex of students; location of school (rural or urban); type of school according to students in attendance (boys only or girls only or co-education); type of school (boarding or day or day/boarding); whilst it is acknowledged that these factors may contribute to differential performance in examinations, they are not considered in detail here because performance in such sub-groups should not be the basis for decision-making for large scale school curriculum and examination planning. The findings which apply to the performance of all candidates should be the basis for such curriculum and examination decision-making. However, the author realises that in looking at the overall performance, an assumption, that the population under investigation is uniform was made.

However, in order to make the sample as representative of the system as a whole as possible a stratified random sample was picked. 'Stratified random sampling utilizes supplementary information about the population in order to draw samples that are more apt to be representatives of that population'(Popham, 1975). Achievement in biology was used as the criterion to stratify the schools into three sub-strata: high achievers; average achievers; and low achievers. The National Examinations Council's rank-order list of school placement for 1981 was used to stratify the schools. The rank-order list shows the comparative performance by schools in each individual subject taken. Thus, the rank-order list for biology was used to stratify the schools for this study.

The rank order list was divided into three (3) sections: top quartile, two middle quartiles, lower quartile, and in order to sample from each group a sample representative of that group, random sampling by discs was used (Popham, 1975). This approach on the sampling gave each school an equal chance of being included in the sample. The school samples picked can therefore be assumed to mirror the range of abilities found nationally, and at the same time provided sufficient numbers of students to allow for appropriate statistical analysis to be undertaken.

Eighty five (85) public secondary schools presented candidates for the 1983 O-level Tanzanian Examination in Biology. Approximately ten percent (10%) of the student population constituted the total pupil sample for this study. The types and number of schools and candidates who were finally included in the sample are shown in Table 4.01 and Table 4.02. As already pointed out, the target population for the pupils' opinion questionnaire for the identification of topics of perceived difficulty were the graduates of 1983 O-level biology course, and the biology teachers who were teaching these students at the time of the main study. Five hundred and nineteen (519), out of the 815, sampled students responded to the Pupils' opinion questionnaire; while forty nine (49) out of 85 fourth year biology teachers returned the opinion questionnaire duly completed. All the fourth year biology teachers (i.e. N=85) were requested to complete the questionnaire which was mailed to the Head of the School.. The 49 returned questionnaire represent over half, (57.6%), of the total expected teacher returns.

Table 4.01

Distribution of sampled schools and candidates according to level of achievement.

Level of Achievement	No. of sampled schools	No. of candidates in the sampled schools	Approximate ratio
High	2	250	1
Middle	4	392	2
Low	2	173	1
Total	8	815	

Table 4.02

Distribution of sampled schools and candidates according to type of school.

Type of school	Boys only	Girls only	Co-Educ.
Boarding	3 N = 296	1 N = 140	
Day	1 N = 123	1 N = 83	
Day/ Boarding	1 N = 73		1 N = 100

4.60 Pilot sample and pilot study.

The pupil opinion questionnaire was piloted twice. First it was administered to a pilot group in Tanzania in July 1983. The pilot group comprised fourth form candidates of one of the public schools which presented candidates for the O-level biology examination in November 1983. This pilot group was considered as being as near as possible an ideal representative of the actual group which responded to the final questionnaire in November 1983. Seventy one pupils participated in the first pilot study.

The purpose of the first pilot study was not to analyse the topics for perceived difficulty. The pilot study was carried out with the hope of uncovering any difficulties in the clarity of the response required from the candidates and any other problems. The information obtained from the first pilot study was used to review the instructions. For example, through the pilot work, some administrative features, such as the time required to complete the questionnaire were finalised. 25-30 minutes was viewed by the supervising teacher as essential. Also the need to explain the instructions orally, in addition to the written instructions to the students was deemed necessary.

Secondly, for an estimation of the reliability of the student responses (stability of the responses), the pupil questionnaire was administered twice to the same group of students during the period of the main study in October/November, 1983. The time interval between the two administrations was 25 days. The time interval between the two administrations has a bearing on the scores of the second administration. If the time interval is short the consistency of the results may be inflated by memory of the responses given the first time. If the time interval is too long, errors of measurement due to the instability in testing procedures may be confused with real changes in student perception as a result of learning or other experiences. That is, the scores on a test-retest with references to time interval will be subjected to two kinds of variations: variation in response to the instrument (test/ examination or questionnaire), at a particular moment in time; and variation in the individual from time to time. The optimum time interval should be determined by the use to be made of the results. In the present study, an interval of 25 days between administrations was considered to be appropriate.

Pearson's Product Moment Correlation Coefficient was determined for the students' responses. (Nie et al, 1975). The test-retest scores were correlated to the magnitude of r = +0.8186. By Coefficient Alpha, (Nie et al, 1975), the reliability of the student responses was found to be +0.952 and therefore it can be claimed that the student questionnaire was a reliable instrument.

Although the teachers' instrument was not piloted in this way, the stability of the responses to Section II can be assumed, by inference from the stability of the responses shown by the students. It is likely to be equal to or higher than that of the student questionnaire because the chance of misinterpretation and mixing up aspects of syllabus areas taught would be greater for students than for their teachers.

Section III of the Teacher questionnaire was validated to some extent through a question on the teachers' and biology panel members' questionnaire which had sought their views on whether the list of objectives constituted an adequate summary of the objectives outlined in the Teachers' Guide (1978); and to suggest additional objectives if the list was inadequate.

Although there was this opportunity for the teachers and panel members to add additional objectives to the supplied list, none availed themselves of this opportunity.

The responses to Section Six of the Teachers' questionnaire on continuous assessment were counter-checked through interviews. A total of 19 fourth form biology teachers, who were attending a biology workshop organized by the Biology Panel of the Institute of Education, during the time of the study were interviewed. The 19 teachers constitute approximately 22% of the fourth form biology teacher population. The interviews were individually conducted and recorded, by the author, on an interview schedule which had been specifically prepared for this purpose. (See Appendix IVf).

A comparison of the interview responses with the questionnaire responses of the 19 teachers interviewed showed a similar pattern of responses. (See Section 9.40 and 9.50).

4.70 Data Processing.

4.71 The examination scores.

The main data source for the evidence of achievement of the curriculum objectives and determination of areas of difficulty, are the examination scores and the pupils' opinion questionnaire.
The marks, which constituted the hard data, awarded to each of the sampled candidates on each question and each component of the examination, were transferred to coded sheets from the candidates' scripts to facilitate the preparation of computer data for computer analysis. Only raw marks as awarded on the scripts were taken.

In each section of the examination papers, the marks were coded for each part - question separately and each section was treated as a separate component of the examination. The treatment of each section as a separate component was based on the assumption that questions within a section are more likely to be similar in the abilities that they examine than questions which are from different sections. A detailed data analysis was carried out of the item scores to determine facility (F) values; discrimination (D) values; reliability of questions; sections; and the entire examination; and the validity of the examination.

The objectives of the examination analysis were to find out whether any of the characteristics of the course or the examination were giving rise to relatively high or low performance either for all students or groups of students. The computations performed were:

Mean values of question facility indices (F);

Mean values of question discrimination indices (D);

F and D values for different components of the examination;

Correlations and intercorrelations between different papers, sections, and within sections;

Factor analyses of the Theory and Practical examinations scores as a whole and by girls and boys scores separately.

The details of each of the analyses procedures and findings is presented in separate chapters which follow.

4.72 <u>The opinion questionnaires</u>.

The different questionnaire responses are analysed separately; and in comparison under the relevant chapters. The pupils' opinion questionnaire, and teachers' opinion questionnaire, Section II, responses on topics of perceived difficulty are reported in Chapter 5.

Analysis of teachers' questionnaire Sections III, and IV; biology panel members questionnaire Section I, and II; and examiners' analysis reports are reported in Chapter 7 under validity of examinations. Responses on teachers' questionnaire Section V, and panel members questionnaire Section III are reported in Chapter 8 under construct validity of examinations.

Responses on continuous assessment are dealt with in Chapter 9.

CHAPTER FIVE

IDENTIFICATION OF TOPICS OF DIFFICULTY IN THE TANZANIAN O-LEVEL BIOLOGY CURRICULUM.

5.10 Introduction and Review of Literature.

Many areas of the school science curriculum have been studied from the stand point of the conceptual demands they have placed upon the pupils studying them. Studies which have been reported on learning difficulties in biology are of three types:

1. Studies which have looked into learning problems due to the high conceptual demands of the curriculum (curriculum 'mis-match') (Shayer, 1974; Okeke, 1980).
2. Studies which have looked into specific topic areas in order to determine areas of general misunderstanding of certain biological ideas and concepts (Deadman and Kelly, 1978; Brumby, 1979; Simpson and Arnold, 1982; Longden, 1982; and Bell and Brook, 1984).
3. Studies which have looked into areas of difficulty within specific curricula for the purposes of identifying problem areas, and thereafter planning teaching strategies (Johnstone and Mahmoud, 1980).

The above mentioned studies are reviewed in the following pages.

Shayer, (1974), made a study of the conceptual demands of the Nuffield O-level biology course. Using Piagetian levels of concept development, he analysed the course demands as revealed by the content of the curriculum materials. He found out that the cognitive levels demanded by the course, especially for Year I and Year II, (i.e. 11 and 12 year old pupils), were high. He identified the work on shape; size and movement as being too difficult for Year II (12 year old pupils); also some aspects of evolution were considered difficult.

Shayer's findings have been criticized by other researchers. For example, Ormerod and Duckworth (1975), argue that Shayer's analysis depends on two assumptions:

1. That the hierarchy of levels of mental development postulated by Piaget can be equated with a distinct progression of mental skills in secondary science. Howe (1974), doubts if this is possible.
2. The equating of large numbers of specific content areas in all three sciences with hypothesized levels of development is also speculative because each varies in inherent difficulty.

Gaskell, further argues that: 'It is a revealing and somewhat difficult experience to decide the levels of conceptual difficulty at which a concept can be taught and then to decide at which Piagetian level of working each of these approaches operate for the full understanding at that level' (Gaskell, 1973, p.374.).

However, Shayer's later work, summarized in 'Towards a Science of Science Teaching', has involved the development of reliable procedures to test the cognitive levels of pupils and has gone some way towards meeting this criticism. (Shayer and Adey, 1981, pp. 26, 61, 62).

Okeke and Wood-Robinson (1980), used interviews to study pupils' understanding of selected biological concepts in Nigeria. The concepts studied were related to growth; reproduction; and transport mechanisms within living organisms. Okeke was interested in the performance of pupils as revealed by scores obtained on an interview schedule, in the form of an oral examination. The pupils' scores awarded by judgemental analysis of the interview responses, were then related to the Piagetian levels of the students' concept development.

From the conceptual analysis, the following were reported as areas of common misconception within the selected topics: asexual reproduction; diffusion; osmosis; cell enlargement and cell differentiation; effects of pressure on transport.

Okeke concluded by pointing out that for teaching to be successful and meaningful, it is always useful to ascertain what the learners already know about the concept to be studied.

Using open-ended interviews, Deadman and Kelly (1978) investigated what school boys understand about evolution and heredity before they are taught these topics. The sample consisted of boys following the Nuffield O-level biology course years 1 - 4. The interviews were organized to cover six concepts: Evolution as a phenomenon; why evolution occurred; the process of change; adaptation; chance; and inheritance. They reported that the concept of adaptation was found to be well established; those of chance and inheritance were least established. The study points out two areas of potential difficulties: the 'naturalistic and Lamarkian' interpretation of most of the concepts; and inadequate understanding of probability.

The concept of 'natural selection' is central to the understanding of evolution and heredity. Difficulties, and misconceptions in understanding the concept of natural selection, have been shown not to be confined to O-level students of biology.

Brumby (1979), reported misunderstandings and difficulties experienced by post-A-level students undertaking further biological studies. Using a written test to investigate the level of understanding of this concept, she reported that only 18 per cent of a group of first year university students with an advanced -level biology background were consistently able to apply the concept to common environmental problems. For example, she points out that, over half of the students mistakenly formulated a 'theory of adaptation by induced mutation', instead of a 'theory of evolution by natural selection.'

Longden (1982), scrutinized major sources of misconceptions and learning difficulties encountered by school pupils in learning genetics. He examined the problem through recorded interviews. He found out that some misconceptions were related to the nature of the concepts used in teaching genetics, such as representation of meiosis by fixed 'inanimate' stage diagrams. Other difficulties he points out are due to separation in teaching time between meiosis and genetics per se; and lack of practical experience for students.

Simpson and Arnold (1982), used interviews to study the 'availability of prerequisite concepts for learning biology at certificate level' in Scotland. They investigated the level of understanding of prerequisite concepts essential for an understanding of photosynthesis regardless of the route taken towards the final concept.

They hypothesized that 'mature, stable, accurate concepts of gas, food, energy, energy conversions, plants, chemical change, cells and carbohydrates' would constitute the necessary 'appropriate cognitive preparation' (p.65). They chose to investigate four of these concepts: living things; gases; energy; and food, using pupils of Primary (aged 11-12 years) and Secondary education, (aged 12-13 years and 13-14 years). The results as regards pupils studying biology at certificate level were that there was failure to classify correctly items into living and non-living; to relate reproduction, respiration, excretion, and sensitivity as characteristics of living things; in the meaning of 'food' as applied to plants and animals; and that starch supplied energy; and in classifying forms of energy (Simpson and Arnold, 1982).

They attributed their findings to: lack of learning opportunities; conceptual demands of concepts on the part of the learner; syllabus demands on the teacher; the number, structure and timing of some of the learning experiences given to the pupils for the development of the concepts, and lack of clear definitions and agreed meaning of terminology.

Bell and Brook (1984), have studied some aspects of secondary students' understanding of plant nutrition. The study aimed at finding out the extent to which the students understood: that plants carry out autotrophic nutrition and not heterotrophic nutrition as do animals; the role of light, chlorophyll and raw materials in photosynthesis; and how the students conceptualised the relationship between food, energy and the maintenance of plant metabolism. They used both written tests and interviews to collect the data.

Bell and Brook reported that only a fifth to a third of 15 year old students used scientifically accepted ideas on plant nutrition in written responses. Some of the ideas used by the students in their written responses include: plants obtain their food from the environment, rather than manufacture it internally; 'food' for plants is anything taken in from the outside, for example water, minerals and air.

They further reported, as a result of the interviews, that there appeared to be little understanding of the relationship of the process of photosynthesis with other physical and chemical processes carried out by plants, for example, water up-take and respiration.

Bell and Brook finally point out that, since teaching involves helping students to construct for themselves the accepted ideas, the starting point for a teaching sequence should be the intuitive ideas students have on the particular concepts to be studied.

As a preliminary step to a much larger study involving the evaluation of the school biology curriculum in Scotland, Johnstone and Mahmoud (1980), determined the relative difficulty with which different topics were perceived within the biology syllabus of the Scottish Certificate of Education Examination Board (SCEEB - 1974).

Johnstone and Mahmoud applied a 'direct approach' method to collect data from pupils, teachers and university students using an opinion questionnaire in which were listed the major topics from the syllabus. Other data sources included examiners' reports, and discussion with inspectors of schools.

Data analysis comprised the computation of facility values of the opinions expressed by the respondents. An index of topic difficulty was calculated using the formula:

$$\text{Index of Difficulty} = \frac{\text{Number of students who said it was difficult X 100}}{\text{Total No. of students} - \text{No. who said the topic was not taught}}$$

(Johnstone and Mahmoud, 1980)

They observed (F) values ranging from 6% to 54%. The topics pointed out as difficult were: chemistry of energy; ATP and ADP; chemistry of photosynthesis; chemistry of respiration; osmosis and water potential; water balance problems and osmo-regulation in living organisms; genes; and the mechanism of evolution (Johnstone and Mahmoud, 1980, p.165).

Johnstone and Mahmoud point out that 'the difficulty of a topic as perceived by pupils will be a major factor in their ability and willingness to learn it' (p.166). And because of this they suggest that where there is a disparity between pupils' and teachers' views, the pupils' views is probably the more important one for the researcher and for the curriculum planner. However, the outstanding point about this work is the degree of agreement between school , university students', and teachers' opinions about the relative difficulty of the topics . In later studies, they sought to find out why the difficulties existed, and how to solve them. (Johnstone and Mahmoud, 1980, 1981).

As in the last study above, the study that is reported in the rest of this Chapter is based on an analysis of a specific curriculum, that is, the Tanzanian O-level Biology Curriculum.

A detailed study of the curriculum topics and sub-topics has been carried out so as to identify areas which are perceived as difficult by students and teachers, and to compare students' and teachers' areas of perceived difficulty using the student and teacher opinion questionnaire responses.

In addition, a detailed analysis of the 1983 O-level biology Examination has been carried out with the intention of identifying topics of 'real' difficulty. A comparison can then be made by matching perception of difficulty with performance in the examination.

The study proceeds by first looking into the purposes and process of examination item analysis.

5.20 Identification of topics of difficulty by item analysis of an examination.

5.21 Item analysis - definition and purposes.

'Item analysis is defined as the process of evaluating the quality of individual examination questions on the basis of the response given to them generally by a representative sample of candidates'. (Willmott and Hall, 1975, p.41).

The purposes of item analysis are many. They are outlined in most educational measurement, and assessment in education texts, for example, Thorndike and Hagen (1969); Eggleston and Kerr (1969); Cronbach (1970); Gronlund (1971); Nuttall and Willmott (1972); Ebel (1972); Ward (1980) and Popham (1981).

For the teacher, information obtained from item analysis can be used as a basis for preparation and presentation of learning experiences. Easily achieved items may indicate areas within the curriculum which have been well taught and learned. Such areas may be omitted from or treated lightly in class discussions. Low performance may indicate unattended areas in curriculum and thus call for attention. Therefore, item analysis provides information about students' abilities and enables the teacher to know the needs of students and where to start or continue at the level of students' capacities by modifying the teaching-learning situation accordingly. When used in formative evaluation item analysis can form part of feedback mechanism to provide insights which lead to the preparation of better tests for classroom use on future occasions by teachers.

For the student, item analysis provides feedback data helping them to improve their learning.

For examiners, item analysis serves two purposes. Firstly, when used in a pre-test situation it helps to detect specific technical faults, such as, ambiguities, clues, ineffective distractors, and other technical defects that went along with the preparation of the test, and thus is a powerful tool for providing information for improving items. Secondly, it provides information on the quality of items. Such information can be used for the selection of items for future use.

For the curriculum planner, item analysis provides feedback information which can be used for curriculum modification if there be need. Item analysis helps in evaluating the appropriateness of specific learning outcomes and the course content. Material that is too difficult for the pupils might suggest curriculum revision or shifts in teaching emphasis. Similarly, errors in pupils' thinking which persistently appear in item analysis data might direct attention to the need for more effective teaching procedures which should be of concern to curriculum planners. Item analysis data thus provide insights into instructional weakness and provide clues for improvement.

5.22 The process of Item analysis.

There are several indices of item analysis, for example, the mean, range, standard deviation, facility and discrimination indices. Other indices include choice index, and the mean ability index. Two of the above indices are of particular interest to the classroom teacher, examiner, and the curriculum planners: these are the facility (F), and the discrimination (D) indices.

The facility index.

This describes the degree of easiness or difficulty of an item. It is the ratio between those candidates who respond correctly to an item to the total number of candidates who attempt the item. This ratio is often expressed as a fraction or the percentage of candidates who get the item right.

What is regarded as an acceptable average facility value and the spread of scores of items is dependent in large part upon the purpose of the test. For example, when the function of a test is to assess mastery, the tester would look for the highest possible values, showing that the students have mastered the basic essentials in the area covered by the test.

When the purpose of the test is to indicate and discriminate between different levels of achievement of different candidates, thus serving as a basis for ranking or grading, the test should then yield a spread of scores. A high degree of discrimination is most likely when questions or items, have a facility value equal to 0.50 (Nuttall and Willmott, 1972).

Facility value will vary for diagnostic tests depending on whether the test is administered at the beginning of instruction in a new area (pre-test), or at the end of instruction (post-test). For the pre-test low facility values can be expected, and mostly high facility values for the post-test. Mastery tests, and diagnostic tests are not concerned with spreading scores.

Discrimination index.

Tuckman describes the discriminability of an item as the 'extent to which a test item is responded to correctly by those students possessing more of the quality being measured and incorrectly by those students possessing less of this quality' (Tuckman, 1972, p.154). The discrimination index thus shows the discrimination ability of the test as it discriminates between candidates with high and low achievement.

Gronlund, (1971), suggests the formula given below to estimate item discrimination index:

$$\text{Discrimination power} = \frac{Ru - Rl}{\frac{1}{2}T}$$

Where: Ru = Number of pupils in the 'upper' group who got the item right.

Rl = Number of pupils in the corresponding 'lower' group who got the item right.

T = Total number of pupils included in the analysis.

There are conditions which govern the use of the above formula, among the most important of which are that the group size must be large, and the candidates should be allowed adequate time to attempt all the items.

There are other ways of calculating the discrimination index, for example, by calculating the biserial correlation, or the point biserial correlation.

The biserial and point biserial correlations involve the use of a computer. Short-cuts to the calculations have been proposed by Connaughton and Skurnik (1969), and Nuttall and Skurnik (1969).

The main point to remember is that the discrimination index, in this respect, is the 'correlation between the item score and the total score on the test' (Nuttall and Willmott, 1972, p.20). Popham emphasises this by saying that it is the correlation coefficient between 'the continuous variable of total test score and the dichotomous variable of performance on a particular item' (Popham, 1981, p.295).

In curriculum evaluation, the discrimination index is not as important as the facility value because the purpose of teaching is not to discriminate between different levels of student achievement. The main interest in curriculum evaluation is to find out what proportion of pupils have grasped a fundamental concept as a result of their course.

However, as in many public examinations, the examinations in Tanzania are less looked upon as tools for curriculum evaluation than as a basis for selection for further educational opportunities; and careers placement. In this study, both facility and discrimination indices of the examination items have been computed in order to provide information about the effectiveness of the curriculum and the efficiency of the examinations.

5.23 Item analysis for examinations in which candidates answer all the items.

Two kinds of analysis can be carried out:

1. Analysis of objective items.
2. Analysis of subjective items.

Analysis of objective items:

Objective items are items 'which are asked in such a way that for each of them there is only one predetermined correct answer'. (Macintosh et.al. 1969, p.9). Multiple choice, true-false, and matching items are examples of objective items. In these kind of items, the responses are either correct or incorrect. Facility values for such items can be calculated as the ratio between those candidates who respond correctly to the total number of candidates in the analysis. This can be expressed as a percentage:

$$\text{Facility (F) value} = \frac{R}{N} \times 100$$

Where: R = Number of candidates with right responses;

N = Total number of candidates in the analysis.

Analysis of subjective items:

With subjective items, 'what is to be accepted as the correct answer is a matter of opinion' (Thyne, 1974, p.177), a matter of opinion with respect to the examiners who may differ in their interpretation and opinion towards question responses, thus showing differences in degrees of subjectivity which could affect the leniency or severity with which the responses are marked. Subjective items include short-answer questions, and essay (extended) answer questions.

Drake defines the facility index of such questions as, 'the sum total of the marks obtained for this question, divided by the total marks for the question and expressed as a percentage' (Drake, 1941, p.392; Nuttall and Willmott, 1972, p.22). The above definition can be expressed by the following formula:

$$F = \frac{\text{Sum of marks obtained by all candidates on the question}}{\text{Sum of maximum marks obtainable by all candidates on that question}}$$

The above formula can be reduced to:

$$F = \frac{\bar{X}_i}{T_i}$$ (Willmott and Hall, 1975, p.45).

Where: $\bar{X}_i$ = Mean score of the item;

T_i = Maximum score possible for that item.

The above formula is useful in that it is simply a generalization of that used with objective items. The interpretation of the index is more difficult with essay questions. A low facility index for an essay question may indicate that it was rather difficult, (the inherent difficulty of the question), or was severely marked. Thus the facility index of subjective questions is susceptible to variations in the leniency or severity of the marker. It is worth noting also that the (F) value of questions in general depends on the nature of the particular group of candidates attempting the question. That is, the mean ability index of the group attempting the question.

5.24 Item analysis for examination questions where candidates have a choice.

Question choice is most common in examinations consisting of essay questions. The most significant feature of essay questions is the freedom of response allowed to the candidates and the fact that no one single answer can be listed as correct and complete and given to clerks to check, but even an expert cannot usually classify a response as categorically right or wrong. Rather there are different degrees of quality or merit which can be recognized (Willmott and Hall, 1975). 'Furthermore, the grades assigned to the papers are usually based in part on what the student has written and in part on what the examiner believes the student meant by what he wrote' (ibid, p.5.).

In discussing issues arising from allowing candidates a choice of questions, Thorndike et.al. (1969), point out that a well prepared essay question requires the student to:

1. Determine what the problem is that is being presented,
2. Review the body of knowledge that he has and select those facts, principles, or ideas that are relevant to the problem,
3. Relate these to one another and organise them into a coherent whole,
4. Produce his answer.

To perform the above tasks, teachers and examiners believe that different levels of thinking are involved. This is one of the reasons why essay type questions are included in most school examinations.

Several reasons are advanced for allowing choice of questions in essay type question papers. One common reason is that, during instruction, the teacher covers broad fields which cannot all be evaluated by a single examination. To allow choice of questions a large area of the curriculum can be covered. On the other hand, to allow choice of questions is to allow candidates to select the question they consider will allow them to display their particular talents to the best advantage. It has been observed that the more able candidates select the easier questions, which in general are questions with a larger content of knowledge (Willmott and Hall, 1975).

Therefore, a spread in the choice of question made by the candidates may indicate instructional coverage over the broad spectrum of the intended curriculum. On the other hand, any unselected items may indicate areas uncovered or insufficiently covered during instruction. Poor performance may be taken as an indication of areas of difficulty, while good performance may be taken as an indication of achievement of curriculum objectives.

Methods for the calculation of facility values and other indices for free-response, free-choice items have been discussed intensively by Morrison (1972); Nuttall and Willmott (1972); Backhouse (1972); and Willmott and Nuttall (1975). They suggest the calculation of four types of indices. These are:-

The choice index (C): This is a measure of the popularity of the question. It is defined as the percentage of the total entry which attempt the question.

The mean ability index (M_T): This is a measure of the ability of the group of candidates attempting the question. It is given by the mean of their total percentage marks.

The facility index (F): This is a measure of the easiness or difficulty of the question. It is defined as the mean percentage mark which a homogenous group of the average ability candidate (M_T = 50 percent) would be expected to obtain on the question.

The discrimination index (D): This is a measure of how the question discriminates between candidates of different abilities and of how it is matched to other questions on the paper. It is defined by the correlation coefficient (r) between the marks on the question for those candidates attempting it and their total mark on the paper minus their marks on the question. That is, the correlation between X_Q and $X_T - X_Q$,

where: X_Q = The mark on the question;

X_T = The total mark on the paper for a particular candidate.

(Nuttall and Willmott, 1972, p.66).

As already described in Chapter Four, the Tanzanian O-level Biology examination papers consist of different types of objective and subjective items. Also it was pointed out that Section A of the theory paper consists of essay questions on which choice is offered to candidates.

The item indices discussed in the previous section are computed on the different items of the Tanzanian biology examination as is applicable. The details of the analysis form the focus of the next section of the present chapter where the procedures taken are described in detail for each analysis.

Further analysis of the examinations' reliability and validity are reported in Chapter 6 and 7 respectively.

5.30 Analysis of the 1983 Tanzanian O-level Biology Examination and the O-level biology curriculum topics.

5.31 Part One: Analysis of the Examination.

A detailed analysis was carried out of the 1983 Tanzanian O-level Biology

Examinations. (See Appendix IIIa and Appendix IIIb). The scripts of 519 candidates were involved in the analysis reported in this chapter.

The objectives of the detailed analysis were to find out whether any of the characteristics of the examinations were giving rise to relatively high or low performance either for all students or groups of students. The performance levels were then related to the curriculum topics being examined with the intention of pointing out course topics which were producing difficulties for all or some students.

Thus with regard to ease or difficulty of the examinations and the curriculum topics, the following studies have been carried out:

1. A study of the individual item characteristics. Two characteristics have been studied for each question, and part question: the Facility (F) index, and the Discrimination (D) index. The detailed item analysis was carried out for each question and part question.

2. A comparative study of the performance on the different components (sections) of the examination. First, the performance on the Theory and Practical examinations are compared. Secondly, the performance, in terms of facility and discrimination powers, on the different sections on each paper are compared: Essay questions section; objective questions section; Short-answer questions section of the Theory examination; and the Experimental investigation section with the Observational investigation section of the Practical examination.

The comparisons have been done by computing means, F-values, D-values, standard deviations, and correlation coefficients between the sections under consideration and the entire examination (Theory + Practical).

3. A comparative analysis of perceived ease or difficulty of the curriculum topics as expressed by students and teachers has also been studied by comparing mean frequency distribution of responses on the opinion questionnaire (see Appendices IVa and IVb), and finally calculation of degree of agreement on the ease or difficulty of the different curriculum topics as scored on the examinations.

Studies on other characteristics of the examination, that is, reliability and validity, which may affect performance, are dealt with in the next two chapters.

Table 5.01 and Table 5.02 show the distribution of marks within the sampled 519 candidate scripts for the Theory and Practical examinations respectively. The 519 scripts are the scripts of all the 519 students who responded to the student questionnaire which has been analysed in Part Two of this Chapter to identify topics of perceived difficulty. Fig. 5.01; Fig. 5.02; and Fig. 5.03 show the relationship between the raw scores and Cumulative Percentage for the Theory examination, Practical examination and a comparison of the two respectively.

5.32 Data analysis and results.

5.33 Facility and Discrimination Powers of the Examination: Theory paper - Section A.

This is the only section, in the entire examination, on which choice was offered to the candidates. It consists of three essay questions. Candidates were required to answer one of the three questions.

To analyse this section, Morrison's approach for analysing choice type questions is used (Morrison 1972). Four separate indices: the choice index (C); the mean ability index (M_T); the facility index (F); and the discrimination index (D); for each of the three questions have been computed. They are presented in Table 5.03. The formulae used are given below:

$$F = \frac{M_T}{\text{Maximum possible score on the question}}$$

D = The correlation between X_Q and $X_T - X_Q$.

Where M_T = Mean of the total mark of candidates attempting the question.

X_Q = Mark on the question

X_T = Total mark on the paper for a particular candidate.

Table 5.01

Distribution of Raw Scores within the Theory Examination.

Raw Score	Absolute Frequency	Cumulative Frequency
5	1	.2
6	1	.4
7	3	1.0
8	11	3.1
9	19	6.7
10	14	9.4
11	19	13.1
12	23	17.9
13	18	21.4
14	33	27.7
15	25	32.6
16	17	35.8
17	23	40.3
18	26	45.3
19	24	49.9
20	28	55.3
21	17	58.6
22	19	62.2
23	18	65.7
24	22	69.9
25	12	72.3
26	21	76.3
27	16	79.4
28	22	83.6
29	12	85.9
30	12	88.2
31	7	89.8
32	7	91.1
33	7	92.5
34	5	93.4
35	6	94.6
36	7	96.0
37	4	96.7
38	5	97.7
39	3	98.3
40	1	98.5
41	3	99.0
42	1	99.2
43	1	99.4
48	2	99.8
53	1	100.0

Maximum Possible Score = 100.000
Mean Score = 20.466
Standard Deviation = 8.252

Sample Size (N) = 519

Fig. 5.01 The relationship between score frequencies and cumulative frequencies (also called cumulative percentage)
The theory examination

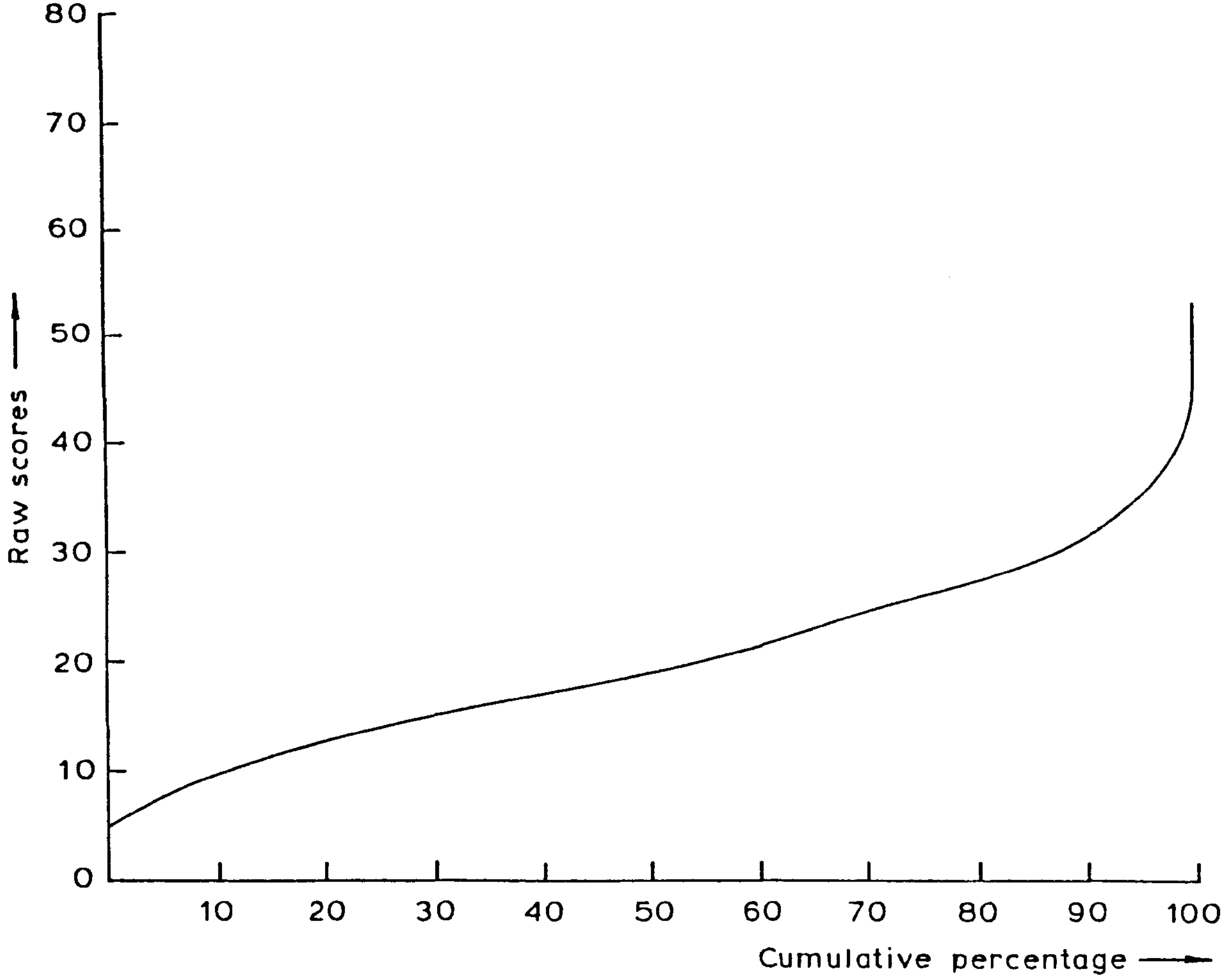


Table 5.02

Distribution of Raw Scores within the Practical Examination.

Raw Score	Absolute Frequency	Cumulative Frequency
0	1	0.2
1	3	0.8
2	9	2.5
3	11	4.6
4	5	5.6
5	8	7.1
6	12	9.4
7	10	11.4
8	10	13.3
9	19	17.0
10	5	17.9
11	16	21.0
12	18	24.5
13	7	25.8
14	16	28.9
15	12	31.2
16	15	34.1
17	15	37.0
18	10	38.9
19	12	41.2
20	13	43.7
21	19	47.4
22	11	49.5
23	12	51.8
24	11	53.9
25	8	55.5
26	18	59.0
27	5	59.9
28	2	60.3
29	11	62.4
30	4	63.2
31	7	64.5
32	8	66.1
33	7	67.4
34	8	69.0
35	8	70.5

Raw Score	Absolute Frequency	Cumulative Frequency
36	7	71.9
37	10	73.8
38	8	75.3
39	13	77.8
40	4	78.6
41	5	79.6
42	6	80.7
43	3	81.3
44	4	82.1
45	5	83.0
46	3	83.6
47	6	84.8
48	8	86.3
49	2	86.7
50	2	87.1
51	5	88.1
52	8	89.6
53	5	90.6
54	9	92.3
55	3	92.9
56	8	94.4
57	3	95.0
58	4	95.8
60	3	96.3
62	2	96.7
63	3	97.3
64	2	97.7
65	3	98.3
66	2	98.7
67	3	99.2
68	2	99.6
73	1	99.8
74	1	100.0

Maximum Possible Score = 100.000
Mean Score = 26.328
Standard deviation = 16.981

Sample Size (N) = 519

Fig. 5.02 The relationship between score frequencies and cumulative frequencies (also called cumulative percentage)

The practical examination

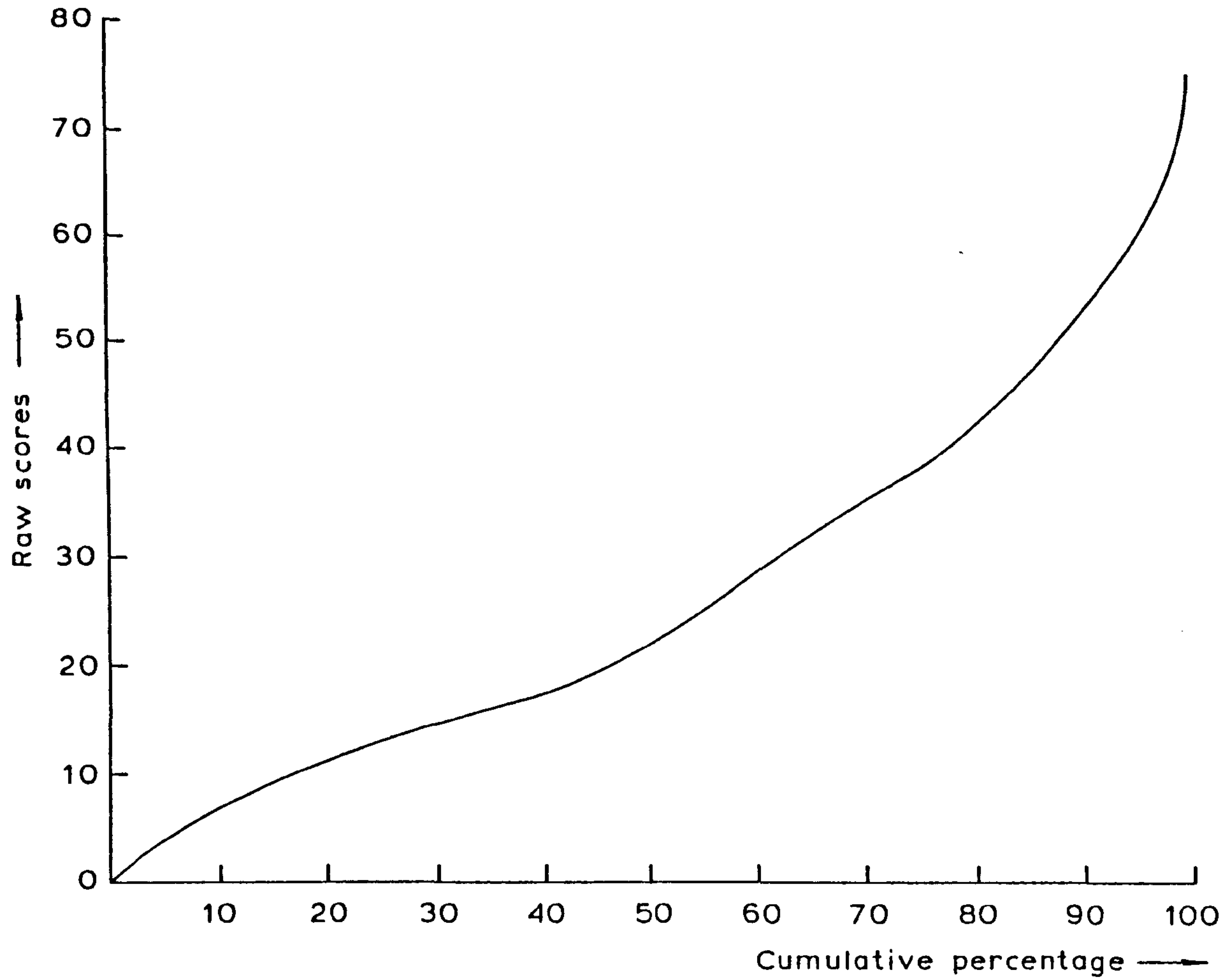
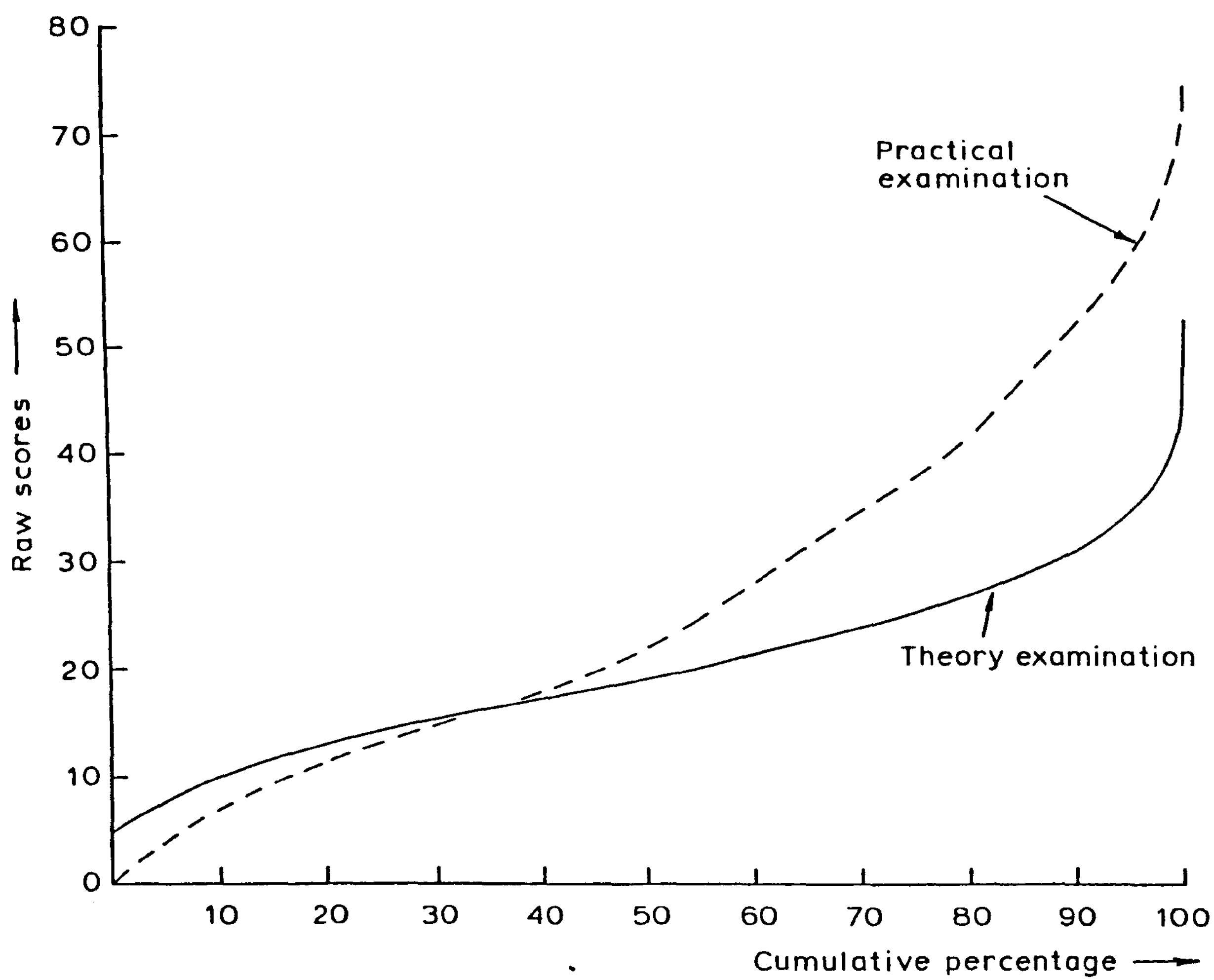


Fig. 5.03 Comparison of the relationship between score frequencies and cumulative frequencies (also called cumulative percentage) for the theory and practical examinations



The Discrimination index given by the point biserial correlation coefficients has been used in a number of other studies of examinations, for example, in an analysis of GCE examinations (Nuttall and Willmott 1972); in studies of the effects of question choice in O-level examinations (Willmott and Hall 1975); in a scrutiny and analysis of objective questions (Duckworth and Hoste 1976); and in an evaluation of a CSE examination in biology (Hoste 1977).

Willmott and Nuttall (1972) have suggested that some care must be taken in choosing the total score with which the item score is to be correlated in calculating a discrimination index through the point biserial approach. They suggest that, in an examination consisting of a number of different parts which are designed to test different facets of attainment, it is usually more appropriate to perform separate item analyses within each part where it may be assumed that each question is designed to test the same facet of attainment within the section.

Therefore, in this study, two kinds of discrimination indices have been computed for each item as follows:

Discrimination index (D)A: Where the item score has been correlated with the total score of the component under consideration.

Discrimination index (D)B: Where the item score has been correlated with the total score of the examination paper under consideration.

The Pearson Correlation Sub-programme of the Statistical Package for the Social Sciences (SPSS) (Nie, Bent and Hull, 1975), run at the University of Hull was used to compute the D-values and other correlation coefficients reported in the rest of this study. All correlation coefficients shown in this chapter and in the rest of this study carry the positive sign (+) except where otherwise indicated.

The computed question statistics for questions one, two, and three are shown in Table 5.03.

Table 5.03

Item statistics for the essay questions in Section A of the Theory examination.

Question No.	Choice Index	Maximum Possible Score	Mean Ability Index	Standard Deviation	Facility Index (F)	Discrimination Index (D) A	Discrimination Index (D) B
1	.748	20	3.947	2.451	.197	.871	.224
2	.086	20	2.771	0.809	.139	.034	(-) .094
3	.052	20	4.048	0.946	.202	.170	.146

Theory paper - Section B.

This section consists of three objective type questions, namely, a multiple choice question consisting of ten items each item with four alternative choices; a matching question consisting of ten items; and a true-false question consisting of ten items. Candidates were required to answer all items. To analyse this section, separate facility (F) values, and discrimination (D) values for each item have been computed using the formulae:

1. $F = \frac{R}{N}$

Where R = Number of candidates with right responses.

N = Total number of candidates in the analysis.

2. The discrimination indices were calculated through point biserial correlation.

   That is,

   D = 'correlation coefficient between the item score and the total score on the test' (test = Theory examination).

The biserial correlation was preferred to the traditional approach of $D = \frac{R_u - R_l}{\frac{1}{2}N}$ (where R_u = right responses by upper group, and R_l = right responses by lower group), because it gives all the individual scores the chance of being included in the analysis of the (D) values.

The computed question statistics are given in table 5.04, Table 5.05, and

Table 5.06 for the multiple choice items; matching items; and true-false items respectively.

Theory paper-Section C.

This consists of five free-response (non-objective) questions with items and sub-items which vary in number between the questions. Candidates were required to answer all of the five questions, their items and sub-items.

In analysing this section, again separate facility (F) values and discrimination (D) values have been computed for each item and/or sub-item using the method for analysing compulsory non-objective items suggested by Nuttall and Willmott (1972), Willmot and Hall (1975) using the formula:

$$F = \frac{\text{Sum of marks obtained by all candidates on the item}}{\text{Sum of maximum marks obtainable by all candidates on the item}}$$

This formula can also be represented by the ratio of the mean score to the maximum possible score for the item.

$$F = \frac{\bar{X}_i}{T_i}$$

Where $\bar{X}_i$ = Mean score for the item

T_i = Maximum possible score for the item

As for the objective questions, the discrimination index has been calculated through point biserial correlations. That is:

D_A = 'the correlation coefficient between the item's score and the total score on the free-response component; and secondly (D_B) the total Score on the theory paper.

The results are given in Tables 5.07 - 5.11 for questions 7 - 11.

Table 5.04

Item statistics for the multiple choice items in Section B of the Theory examination.

	Item No.	Maximum Possible Score	Mean Score	Standard Deviation	Facility Index (F)	Discrimination Index (D) A	Discrimination Index (D) B
4	i	.5	.363	.223	.726	.392	.172
5	ii	.5	.351	.229	.702	.448	.289
6	iii	.5	.085	.188	.170	.283	.121
7	iv	.5	.138	.224	.276	.329	.147
8	v	.5	.227	.249	.454	.442	.292
9	vi	.5	.229	.249	.458	.508	.369
10	vii	.5	.377	.216	.754	.398	.226
11	viii	.5	.095	.197	.190	.419	.346
12	ix	.5	.274	.249	.548	.523	.521
13	x	.5	.284	.248	.568	.419	.269

Table 5.05

Item statistics for the Matching items in Section B of the Theory examination.

	Item No.	Maximum Possible Score	Mean Score	Standard Deviation	Facility Index (F)	Discrimination Index (D) A	Discrimination Index (D) B
14	i	.5	.300	.245	.600	.524	.397
15	ii	.5	.168	.236	.336	.440	.303
16	iii	.5	.295	.246	.590	.607	.449
17	iv	.5	.086	.189	.172	.368	.209
18	v	.5	.069	.173	.138	.406	.283
19	vi	.5	.151	.230	.302	.571	.490
20	vii	.5	.159	.233	.318	.520	.353
21	viii	.5	.223	.249	.446	.444	.298
22	ix	.5	.089	.191	.178	.440	.394
23	x	.5	.304	.244	.608	.612	.454

Table 5.06

Item statistics for the true-false items in Section B in the Theory examination.

	Item No.	Maximum Possible Score	Mean Score	Standard Deviation	Facility Index (F)	Discrimination Index (D) A	Discrimination Index (D) B
24	i	.5	.070	.174	.140	.216	.080
25	ii	.5	.338	.234	.676	.331	.083
26	iii	.5	.226	.249	.452	.339	.171
27	iv	.5	.282	.248	.564	.339	.020
28	v	.5	.260	.250	.520	.393	.192
29	vi	.5	.359	.225	.718	.238	(-).132
30	vii	.5	.224	.249	.448	.366	.212
31	viii	.5	.322	.240	.644	.441	.209
32	ix	.5	.350	.230	.700	.395	.151
33	x	.5	.387	.209	.774	.307	.224

Table 5.07

Item statistics for the short-answer items in question 7 in Section C of the Theory examination.

	Item No.	Maximum Possible Score	Mean Score	Standard Deviation	Facility Index (F)	Discrimination Index (D) A	Discrimination Index (D) B
34	a	2.0	.233	.307	.117	.567	.559
35	b	0.5	.067	.171	.134	.472	.431
36	c - i	0.5	.175	.239	.350	.551	.504
37	ii	0.5	.135	.222	.270	.519	.466
38	iii	0.5	.095	.197	.190	.553	.535
39	iv	0.5	.258	.438	.516	.566	.520
40	v	0.5	.005	.050	.010	.176	.179

Table 5.08

Item statistics for the short-answer items in question 8 in Section C in the Theory examination.

	Item No.	Maximum Possible Score	Mean Score	Standard Deviation	Facility Index (F)	Discrimination Index (D) A	Discrimination Index (D) B
41	a - i	1.5	.154	.268	.103	.479	.457
42	ii	0.5	.050	.150	.100	.353	.329
43	iii	0.5	.080	.183	.160	.431	.409
44	iv	1.0	.119	.228	.119	.522	.509
45	v	1.0	.178	.366	.178	.433	.426
46	b - i	2.0	.039	.172	.020	.367	.363
47	ii	0.5	.058	.163	.116	.274	.220
48	iii	1.0	.032	.137	.032	.197	.167
49	iv	2.0	.266	.401	.133	.580	.581
50	v	1.5	.103	.253	.069	.389	.393
51	c i	1.0	.050	.180	.050	.326	.305
52	ii	1.0	.041	.169	.041	.298	.267
53	iii	1.0	.089	.250	.089	.251	.218
54	iv	1.0	.153	.262	.153	.358	.340

Table 5.09

Item statistics for the short-answer items in question 9 in Section C in the Theory examination.

	Item No.	Maximum Possible Score	Mean Score	Standard Deviation	Facility Index	Discrimination Index (D) A	Discrimination Index (D) B
55	a - i	2.00	.974	.667	.487	.562	.581
56	ii	1.33	.320	.404	.241	.642	.648
57	iii	1.00	.331	.312	.331	.547	.567
58	iv	1.67	.179	.190	.107	.505	.506
59	b - i	1.67	.228	.357	.137	.543	.558
60	ii	0.33	.026	.090	.079	.326	.301
61	iii	7.67	.098	.311	.013	.430	.427
62	c - i	8.00	.665	.781	.083	.590	.486
63	ii	0.67	.041	.129	.061	.399	.383

Table 5.10

Item statistics for the short-answer items in question 10 in Section C in the Theory examination

	Item No.	Maximum possible Score	Mean Score	Standard Deviation	Facility Index (F)	Discrimination Index (D) A	Discrimination Index (D) B
64	i	1.0	.183	.416	.183	.301	.275
65	ii	1.0	.025	.168	.025	.219	.244
66	iii	1.0	.110	.313	.110	.334	.312
67	iv	1.0	.002	.044	.002	(-).038	(-).056
68	v	1.0	.008	.088	.008	.122	.144
69	vi	1.0	.060	.237	.060	.357	.350
70	vii	1.0	.010	.098	.010	.136	.133
71	viii	1.0	.052	.222	.052	.297	.300

Table 5.11

Item statistics for the short-answer items in question 11 in Section C in the Theory examination.

	Item No.	Maximum possible Score	Mean Score	Standard Deviation	Facility Index (F)	Discrimination Index (D) A	Discrimination Index (D) B
72	a	3.0	.102	.361	.034	.367	.355
73	b	3.0	.004	.062	.001	.146	.146
74	c	4.0	.089	.465	.022	.240	.252
75	d	2.0	.154	.407	.077	.367	.347

The Practical Examination.

The Practical examination consists of three free-response questions. The three questions have been classified initially by the author into two sections: the Experimental investigation section which covers question one; and the Observational investigation section which covers questions two and three. (Appendix IIIb).

The classification is based on the fact that question one requires the candidates to carry out a series of experimental investigations on given specimens; while questions two and three require the candidate to make observational investigations on given specimens. All three questions were compulsory.

To analyse this paper for facility (F) and discrimination (D) indices, the approach used to analyse Section C of the Theory paper has been applied. The results are given in Table 5.12; Table 5.13 and Table 5.14 for questions one, two and three respectively.

5.40 Discussion of the Results.

The Facility and Discrimination Indices

The item analysis reported provides quantitative information about the way each question, item and sub-item contributed to the overall performance of the examination. Considerable variations in the facility and discrimination indices can be seen in both the Theory and Practical examinations.

The Facility Indices: Facility index ranges between 0.30 and 0.80 for objective items and between 0.20 and 0.80 for free response items can be described as satisfactory (Duckworth and Hoste 1976; Mathews1974).
In the Theory examination, the facility indices range from 0.001 to 0.774. The F-indices of the objective questions (multiple choice, matching, and true-false items), are higher than those of the subjective questions (essay

Table 5.12

Item statistics for the free-response items in question I in the Practical examination.

	Item No.	Maximum Possible Score	Mean Score	Standard Deviation	Facility Index (F)	Discrimination Index (D) A	Discrimination Index (D) B
1	a	35.0	10.033	10.367	.287	.983	.937
2	b	4.0	0.674	0.963	.169	.843	.819
3	c	4.0	0.728	1.005	.182	.843	.817
4	d	4.0	0.316	0.628	.079	.580	.577
5	e	6.0	1.056	1.450	.176	.815	.784

Table 5.13

Item statistics for the free-response items in question 2 in the Practical examination.

	Item No.	Maximum Possible Score	Mean Score	Standard Deviation	Facility Index (F)	Discrimination Index (D) A	Discrimination Index (D) B
6	a	3.0	1.289	.839	.430	.413	.362
7	b	1.5	0.774	.750	.516	.616	.499
8	c - i	1.5	0.820	.747	.516	.616	.499
9	ii	1.5	0.451	.695	.301	.464	.277
10	iii	6.0	0.474	.723	.079	.511	.448
11	iv	1.5	0.231	.542	.361	.451	.343

Table 5.14

Item statistics for the free-response items in question 3 in the Practical examination.

	Item No.	Maximum Possible Score	Mean Score	Standard Deviation	Facility Index (F)	Discrimination Index (D) A	Discrimination Index (D) B
12	a - i	7.0	4.025	1.819	.575	.707	.505
13	-ii	15.0	2.184	1.911	.146	.705	.468
14	b - i	1.0	0.643	0.479	.643	.413	.290
15	- ii	1.0	0.308	0.462	.308	.483	.413
16	- iii	2.0	0.127	0.372	.064	.261	.211
17	- iv	2.0	0.418	0.498	.209	.479	.366
18	- v	2.0	0.921	0.291	.461	.383	.242
19	- vi	2.0	0.856	0.352	.176	.430	.345

and short-answer items). Seven out of the 30 objective items have F-values below 0.300, ranging from 0.138 to 0.276. The remaining 23 items have F-value ranging from 0.302 to 0.774. On the other hand, only six out of the 45 subjective questions and items have F-values above 0.200, ranging from 0.241 to 0.516, while the rest, 39 items, have F-values ranging from 0.001 to 0.190.

In the Practical examination, the facility indices range from 0.064 to 0.643. The facility indices for questions of the Observational investigation section are higher than those of the questions of the Experimental investigation section. For all the five items of the experimental investigation section, the F-values lie between 0.079 and 0.287. As in the Theory examination, there is a considerable variation in the F-values for the 14 items of the Observational investigation section. The values for this particular section lie between 0.064 and 0.643. Only four of the 14 items have F-values below 0.200, between 0.064 and 0.176.

The Discrimination Indices.

A Discrimination index, by biserial correlation coefficient, of at least r = 0.20 has been described as satisfactory (Duckworth and Hoste, 1976). Values ranging between r = 0.20 and 0.30 have been described as fair, and values greater than r = 0.30 as satisfactory (Willmott and Nuttall, 1972), while values of 0.60 and 0.70 have been described as acceptable discrimination indices for non-objective items including the traditional O-level essay type questions, (Willmott and Hall, 1975). In both the Theory and the Practical examinations, the discrimination indices (D)A, where the items scores were correlated with the total component score, are generally higher than discrimination indices (D)B, where the items scores were correlated with the total paper score.

Discrimination indices (D)A for the Theory examination range from (-)0.038 to 0.642, while discrimination indices (D)B range from (-)0.094 to 0.648.

The objective items seem to have lower discrimination indices than the free-response items. All the discrimination indices (D)A are above 0.20, while only 10 of the 30 objective items have discrimination indices (D)B below the 0.20 level. On the other hand, the free-response items seem to have a bigger range of the (D)-values with (D)A values from (-)0.038 to 0.642, and (D)B values from (-)0.056 to 0.648. Only 5 out of the 45 items have discrimination indices lower than 0.20.

In the Practical examination, all the D-values are positive. The experimental section has higher discrimination indices than the observational investigation section. The D-values in the experimental investigation section range from (D)A = 0.580 to 0.983, and (D)B = 0.577 to 0.937; while in the observational investigation section the ranges are (D)A = 0.383 to 0.707, and (D)B = 0.242 to 0.505.

<u>The (F) and (D) values for the different examination components.</u>

With regard to paper components, as seen in Table 5.15, and Table 5.16,

Table 5.15

Statistics for the performance on the different components of the Theory examination.

Paper Component	Maximum Possible Score	Mean Score	Standard Deviation	Facility Index (F)	Discrimination Index (D)
Essay	20	3.339	2.323	.167	.630
Multiple	5	2.423	0.951	.485	.671
Matching	5	1.843	1.118	.369	.736
True-False	5	2.820	0.787	.564	.360
Free-Response	65	10.041	5.332	.154	.946
Entire Theory Paper	100	20.466	8.252	.205	

Table 5.16

Statistics for the performance on the different components of the Practical examination.

Paper Component	Maximum Possible Score	Mean Score	Standard Deviation	Facility Index (F)	Discrimination Index (D)
Experimental Investigation	53	12.807	13.398	.242	.957
Observational Investigation	47	13.521	5.702	.288	.730
Entire Practical Paper	100	26.328	16.981	.263	

which show the performance on the different components of the Theory and Practical examinations respectively, the objective questions, (multiple choice, matching and true-false), have higher F-values than the subjective questions, (essay and free-response). The true-false section has the highest F-values of 0.564. On the other hand, and although it is not always the case that the question with low F-values have high D-values, the free-response section in this case has the lowest F-value of 0.154, and highest D-value of 0.946.

The observational investigations section of the Practical examination has a higher F-value than the experimental investigational section, while the D-values of the two sections are 0.957 and 0.730 for the experimental and observational sections respectively. On the whole the Practical examination paper, seem to have a higher F-value than the Theory examination paper.

F-value for the Theory paper = 0.205.

F-value for the Practical paper = 0.263.

Overall, it can be said that both examination papers seem to have low F-values, and high D-values.

5.50 Part Two: Analysis of the Curriculum Topics.

Ease or Difficulty of Curriculum Topics.

In this part of the work, ease or difficulty of the curriculum topics has been approached from two aspects. The first is based on the teachers' and students' perceptions of the level of ease or difficulty of the various topics and subtopics of the biology course as shown by topic/subtopic rating scores on the Teachers' and Students' Opinion Questionnaire. (See Appendix IVa and Appendix IVb).

The second is based on the actual performance by students in the different topics of the course as shown by their performance in the examination questions identified to be assessing the particular topic of the course.

5.51 Analysis of Perceived Topic Difficulty using the Teachers'/Students' Opinion Questionnaire.

The opinion questionnaires for the teachers and pupils have been described in Chapter 4 (and are presented as Appendices IVa and IVb). They consist of a list of topics which constitute the Biology Syllabus, as outlined in the Secondary School Science Syllabuses Volume 2 (1976) (See Appendix I). Each of the topics, and sub-topics is treated as an item in this section of the study. The summary list of the 16 topic areas is given in Table 5.17.

To identify the relative difficulty with which topics were perceived (F) values have been computed on the responses given by both teachers and students. The data in the opinion questionnaire was collected in nominal categories: Easy; Average; Difficult; and Not covered/Not taught. The name of each category was replaced by a number score assigned during data analysis in the order:

Easy = 3
Average = 2
Difficult = 1
Not taught = 0

Table 5.17

List of the 16 Topic areas for the O-level biology course.

Curriculum Topic
1. Classification
2. Cell structure and function
3. Soil and plant growth
4. Nutrition in plants and animals
5. Respiration in plants and animals
6. Transport system in plants and animals
7. Excretion and Osmoregulation in mammals
8. Control and co-ordination systems in organisms
9. Movement
10. Growth and Development
11. Reproduction
12. Genetics
13. Evolution
14. Ecology and interdependence of organisms
15. Parasites and diseases of man
16. Man and his natural resources

Therefore, each respondent was awarded a score on each topic area. Instead of taking the difficulty of a topic as the percentage of only those who said the topic is difficult, (Johnstone and Mahmoud, 1980), in this study it has been assumed that there is no clear cut boundary between easy, average, and difficult. The degree of difficulty for each sub-topic, therefore, has been calculated using the proportion of the mean rating score to the maximum possible rating score for each sub-topic. The mean rating score for a topic equals the average of the rating of the sub-topics.

Therefore, the formula below has been used to compute the (F) values for each topic area:

$$F = \frac{\overline{X}_t}{T_t} = \frac{\text{Mean Topic rating score}}{\text{Maximum possible Topic rating score}}$$

The mean topic rating scores shown in table 5.19 have been calculated using the mean sub-topic rating scores shown in table 5.18.

5.60 Discussion of the Results.

The topics which seem to present both learning and teaching difficulties are: Evolution; Parasites and Diseases of man; Man and his natural resources; Transport system in plants and animals; Ecology and interdependence of organisms; and Genetics. These topics were rated as the most difficult topics in that order, by both teachers and students.

A more detailed examination of the subtopics' (F) values shown in Table 5.18, reveals the areas of perceived difficult by students as being: (the list is given in descending order of difficulty): Man and pathogens - types of pathogens and disease causes; Effects of smoking on health; the relationship between size, surface area and the need for transport system in large organisms; Evidence of evolution - the process of change and natural selection; Chromosomes, genes, mutation - structure of DNA, mitosis and meiosis;

Table 5.18

Teachers' and Students' opinions on Ease or Difficulty of Topic Areas: Mean topic rating: F-values and mean ranking of sub-topics.

Topic	Sub-Topic	Mean rating Score		Facility Index (F)		Mean Ranking	
		Students	Teachers	Students	Teachers	Students	Teachers
1	1	2.599	2.653	.866	.884	1	5
	2	2.170	2.020	.723	.673	17	36
	3	2.182	1.918	.727	.639	15	39
	4	2.173	2.082	.724	.694	16	34
	5	2.583	2.592	.861	.864	2	12
2	6	2.374	2.653	.791	.884	4	5
	7	2.442	2.837	.814	.946	3	1
3	8	2.344	2.429	.781	.809	8	19
	9	1.847	2.429	.616	.809	29	19
	10	2.068	2.224	.689	.741	23	26
4	11	2.222	2.367	.740	.789	13	21
	12	2.373	2.735	.791	.912	5	3
	13	1.992	2.592	.664	.864	26	12
	14	1.835	2.490	.611	.830	31	17
	15	2.248	2.735	.749	.912	12	3
5	16	1.323	2.163	.441	.721	43	29
	17	2.174	2.224	.716	.741	19	26
	18	2.116	2.621	.705	.871	20	9
6	19	1.969	2.571	.656	.857	28	14
	20	1.819	1.633	.606	.544	34	45
	21	2.074	2.347	.691	.782	22	22
	22	1.089	2.633	.363	.878	44	8
7	23	2.332	2.510	.777	.837	9	16
8	24	1.988	1.776	.663	.592	27	42
	25	2.302	2.612	.767	.871	10	9
	26	1.822	2.224	.607	.741	33	26
	27	2.066	2.020	.689	.673	24	36
9	28	1.837	2.531	.612	.844	30	15
	29	2.186	2.265	.729	.755	14	25
	30	1.745	2.061	.582	.687	39	35
	31	2.149	2.122	.716	.707	18	31
10	32	1.788	1.653	.596	.551	36	44
	33	2.350	2.612	.783	.871	7	9
	34	2.278	2.469	.759	.823	11	18
11	35	2.028	1.816	.676	.605	25	41
	36	2.115	2.776	.705	.925	21	2
	37	2.370	2.653	.790	.884	6	5
12	38	1.641	1.714	.547	.571	41	43
	39	1.758	2.143	.586	.714	38	30
	40	1.773	2.286	.590	.762	37	24
13	41	1.344	1.837	.448	.612	42	40
14	42	1.810	2.122	.603	.707	35	31
15	43	1.089	1.959	.363	.653	45	38
	44	1.832	2.306	.611	.769	32	23
16	45	1.740	2.102	.560	.701	40	33

The mean ranking is given in ascending order of difficulty.
That is, No.1 signifies the easiest subtopic and 45 the most difficult.

Table 5.19

Teachers' and students' opinions on Ease or Difficulty of Topic areas: A comparison of Students' and Teachers' topic mean rating score: facility index (F): and mean ranking.

Topic Area	Mean rating score		Facility Index (F)		Ranking of mean rating score	
	Students	Teachers	Students	Teachers	Students	Teachers
Classification	2.510	2.253	.837	.751	2	8
Cell structure and function	2.548	2.745	.849	.915	1	1
Soil and plant growth	2.248	2.361	.749	.787	6	5
Nutrition in plants and animals	2.366	2.584	.789	.861	3	2
Respiration in plants and animals	1.916	2.333	.639	.778	10	6
Transport system in plants and animals	1.801	2.296	.600	.765	13	7
Excretion and Osmoregulation in mammals	2.332	2.510	.777	.837	4	3
Control and co-ordination systems in organisms	2.143	2.158	.714	.719	8	11
Movement	2.124	2.245	.708	.748	9	9
Growth and Development	2.210	2.245	.737	.748	7	9
Reproduction	2.320	2.415	.773	.805	5	4
Genetics	1.898	2.048	.633	.683	11	15
Evolution	1.344	1.837	.448	.612	16	16
Ecology and interdependence or organisms	1.810	2.122	.603	.707	12	13
Parasites and diseases of man	1.610	2.133	.537	.711	15	12
Man and his natural resources	1.740	2.102	.580	.701	14	14

The ranking order is given in ascending order of difficulty
That is, No.1 the easiest and 16 the most difficult.

Conservation of the natural resources - pollution of environment; Types, structure and arrangement of muscles; Mendel's first law of inheritance (dominant and recessive genes, phenotype and genotype); The importance of genetics in plant and animal breeding; and the process of cell division, enlargement and differentiation. The sub-topics listed above have student mean rating scores, and (F) values much lower than the students' grand mean rating score and (F) values of 2.007 and 0.669 respectively.

The extent of agreement between pupils' and teachers' opinions of ease or difficult of the sub-topics was further assessed using the Pearson Chi-Square test Sub-programme 'Crosstabs' of the SPSS (Nie, Bent and Hull, 1975). The pupils' and teachers' agreement on 17 of the 45 sub-topics are significant beyond the level of $p = 0.001$ with X^2 values ranging from 11.75 to 43.56, df = 3. In fact the agreement on 8 out of the 17 sub-topics is significant beyond the level of $p = 0.0001$, with 3 degrees of freedom, that is a perfect agreement. The 17 sub-topics are sub-topics 7;9; 12; 13; 14; 15; 16; 18; 19;22;26; 28;36;41;42;43; and 44. The level of agreement on a further 10 sub-topics was $p = 0.01$ with X^2 values ranging from 7.9 to 11.27, df = 3. The 17 sub-topics are identifiable from Appendix IVa, and Appendix IVb: Section II.

The mean topic facility indices for teachers and pupils were ranked in increasing order of difficulty. A further comparison of the students' and teachers' opinions on the topics was then obtained by computing correlation coefficients between students' and teachers' mean topic scores; mean topic facility indices; and mean ranking of topic facility indices.
The results shown in Table 5.20 further indicate the extent to which students and teachers agree on their opinions, the respective correlation coefficients being 0.8035; 0.8031 and 0.8069.

Finally, the pattern of agreement revealed by the data is shown in

Fig. 5.04. There is a high agreement between students' and teachers' opinions. The topics identified as being the most difficult by the students are the same ones that teachers have identified as being difficult.

However, from Table 5.19, two obvious mismatches between teachers' and pupils' views were reflected in topic No.1 and topic No.6. Topic No.1 (classification), and topic No.6 (transport system in plants and animals), have been ranked 2nd and 13th by students, while teachers have ranked them as 8th and 7th respectively.

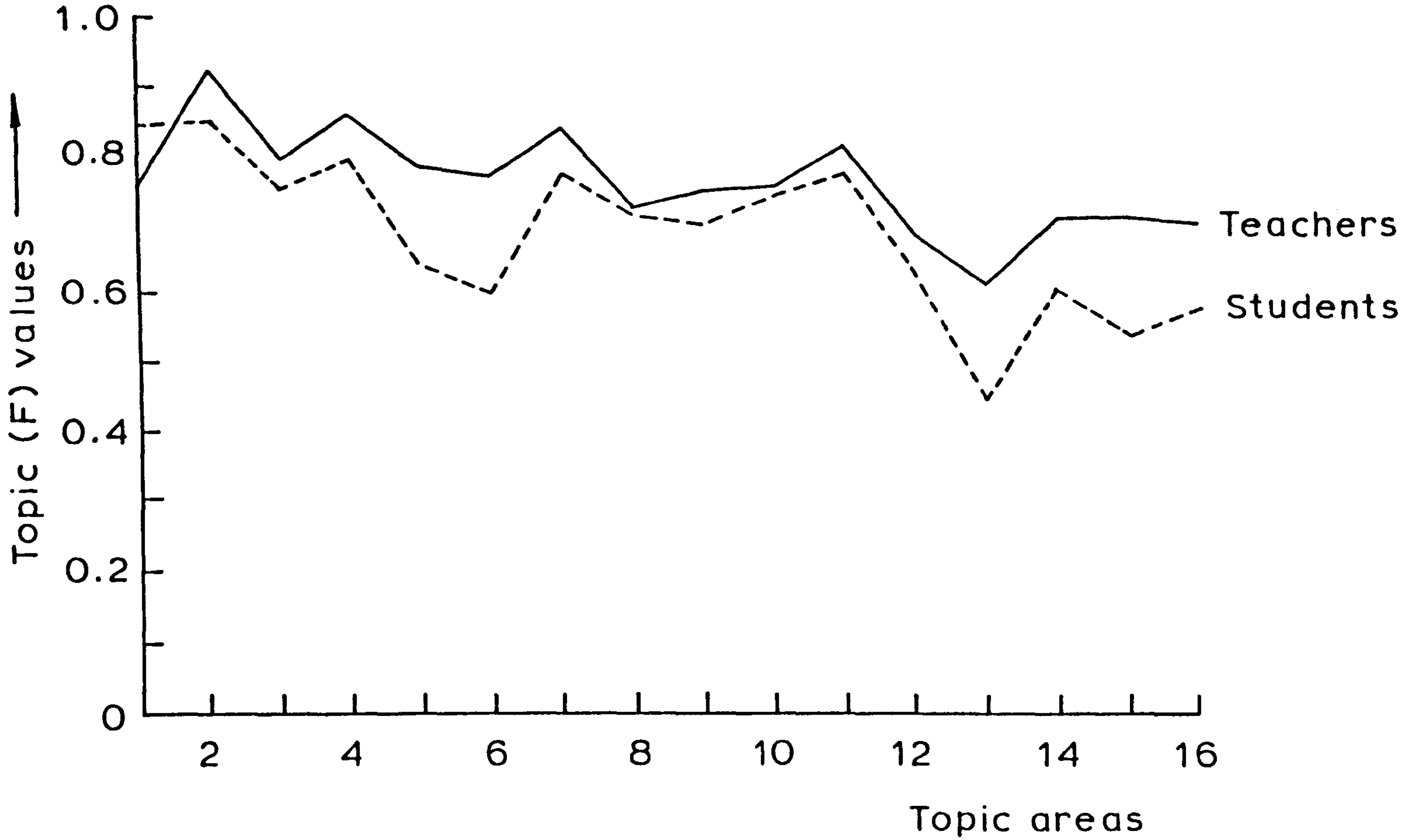
Table 5.20

A comparison of students' and teachers' perceived difficulty through the correlation between students' and teachers' mean topic scores; mean topic facility indices; and mean ranking of topic facility indices.

	Teachers' mean topic Scores	Teachers' mean topic facility indices	Teachers' mean ranking of topic facility indices
Students' mean topic Scores	0.8035		
Students' mean topic facility indices		0.8031	
Students' mean ranking of topic facility indices			0.8069

It has been suggested that, 'the difficulty of a topic as perceived by pupils will be a major factor in their ability and willingness to learn it' (Johnstone and Mahmoud, 1980). And because of this they suggest that where there is a disparity between pupils' and teachers' views, the pupils' views is probably the more important one for the researcher and for the curriculum

Fig. 5.04 Relationship between Students' and Teachers' perceived ease or difficulty of topic areas



planner. This point on differences between pupils' and teachers' views is further discussed in Section 5.71; Section 8.83 and Section 10.10.

5.70 Content Analysis.

An analysis of the topics tested by the examination questions was carried out using an examiners' classification of the questions and the actual examination scores obtained on them.

The content of the syllabus, as already pointed out in Chapter 4, was grouped into 16 main topic areas with 45 sub-topic areas altogether. Grouping of the items on the exam papers by topics was performed to obtain the aggregate performance by the candidates on all the items which have been identified and grouped together as belonging to the same topic content category. The grouping of the items was done by 8 experienced examiners. The procedures and materials used by the examiners have been described in Chapter 4, (also see Appendix IVd.) The extent of agreement between the 8 examiners is discussed in Section 7.53.

Having allocated each question, item and sub-item to its respective topic area, the next step was to calculate a topic score for each of the 16 topic areas for each candidate. This was achieved by adding together the scores in all the items associated with the particular topic under consideration for each candidate.

The topic scores were then used to compute topic means and facility indices for each topic area. In addition, the perceived topic facility indices discussed in the previous section were intercorrelated with the actual facility indices, to determine the relationship between the perceived and actual topic difficulty: The Pearson Correlation Sub-Programme SPSS (Nie, Bent and Hull, 1975) was used throughout this analysis. The results are shown in Table 5.21; and Table 5.22; for the Theory and Practical examinations respectively; and Table 5.23 shows the comparison of (F) values for perceived and actual difficulty, while Table 5.24 showsthe intercorrelation coefficients between actual and perceived difficulty.

Table 5.21

The distribution of Questions and score statistics by Topic areas in the Theory examination.

	Topic Area	Total No.of Questions	Maximum Possible Score	Mean Score	Facility Index (F)
1	Classification	11	6.50	2.261	.349
2	Cell structure and function	5	2.50	0.955	.382
3	Soil and plant growth	4	21.50	3.536	.164
4	Nutrition in plants and animals	15	13.50	2.730	.202
5	Respiration in plants and animals	9	21.50	1.325	.062
6	Transport system in plants and animals	15	26.67	7.748	.291
7	Excretion and Osmoregulation in mammals	6	7.50	1.121	.149
8	Control and co-ordination systems in organisms	8	8.50	1.076	.127
9	Movement	2	1.00	0.487	.487
10	Growth and Development	7	6.00	1.565	.261
11	Reproduction	6	24.00	1.133	.047
12	Genetics	5	22.50	0.899	.040
13	Evolution	5	23.00	0.925	.040
14	Ecology and interdependence of organisms	6	22.67	3.773	.166
15	Parasites and diseases of man	4	22.50	0.772	.034
16	Man and his natural resources	2	20.00	3.175	.159

Table 5.22

The distribution of questions and score statistics by Topic areas in the Practical examination.

	Topic Area	Total No.of Questions	Maximum Possible Score	Mean Score	Facility Index (F)
1	Classification	4	11.5	3.226	.281
2	Cell structure and function	1	6.0	0.474	.079
3	Soil and plant growth	-	-	-	-
4	Nutrition in plants and animals	16	91.0	23.561	.259
5	Respiration in plants and animals	-	-	-	-
6	Transport system in plants and animals	2	22.0	6.209	.282
7	Excretion and Osmoregulation in mammals	2	4.5	2.063	.458
8	Control and co-ordination systems in organisms	-	-	-	-
9	Movement	-	-	-	-
10	Growth and Development	1	2.0	0.418	.209
11	Reproduction	6	10.0	3.274	.327
12	Genetics	-	-	-	-
13	Evolution	-	-	-	-
14	Ecology and interdependence of organisms	2	3.0	0.682	.227
15	Parasites and diseases of man	-	-	-	-
16	Man and his natural resources	-	-	-	-

Table 5.23

Comparison of student's and teacher's perceived topic facility indices (F) with student's actual mean topic facility indices (F) for the entire examination (Practical + Theory).

Curriculum Topic	Preceived Topic Facility Index (F)		Actual Mean Topic Facility Index (F)
	Student	Teacher	
Classification	.837	.751	.305
Cell structure and function	.849	.915	.168
Soil and plant growth	.749	.787	.164
Nutrition in plants and animals	.789	.861	.252
Respiration in plants and animals	.639	.778	.062
Transport system in plants and animals	.600	.765	.290
Excretion and Osmoregulation in mammals	.777	.837	.265
Control and co-ordination systems in organisms	.714	.719	.127
Movement	.708	.748	.487
Growth and Development	.737	.748	.248
Reproduction	.773	.805	.130
Genetics	.633	.683	.040
Evolution	.448	.612	.040
Ecology and interdependence of organisms	.603	.707	.174
Parasites and diseases of man	.537	.711	.034
Man and his natural resources	.580	.701	.159

Table 5.24

Comparison of Perceived and Actual Topic Difficulty by intercorrelations.

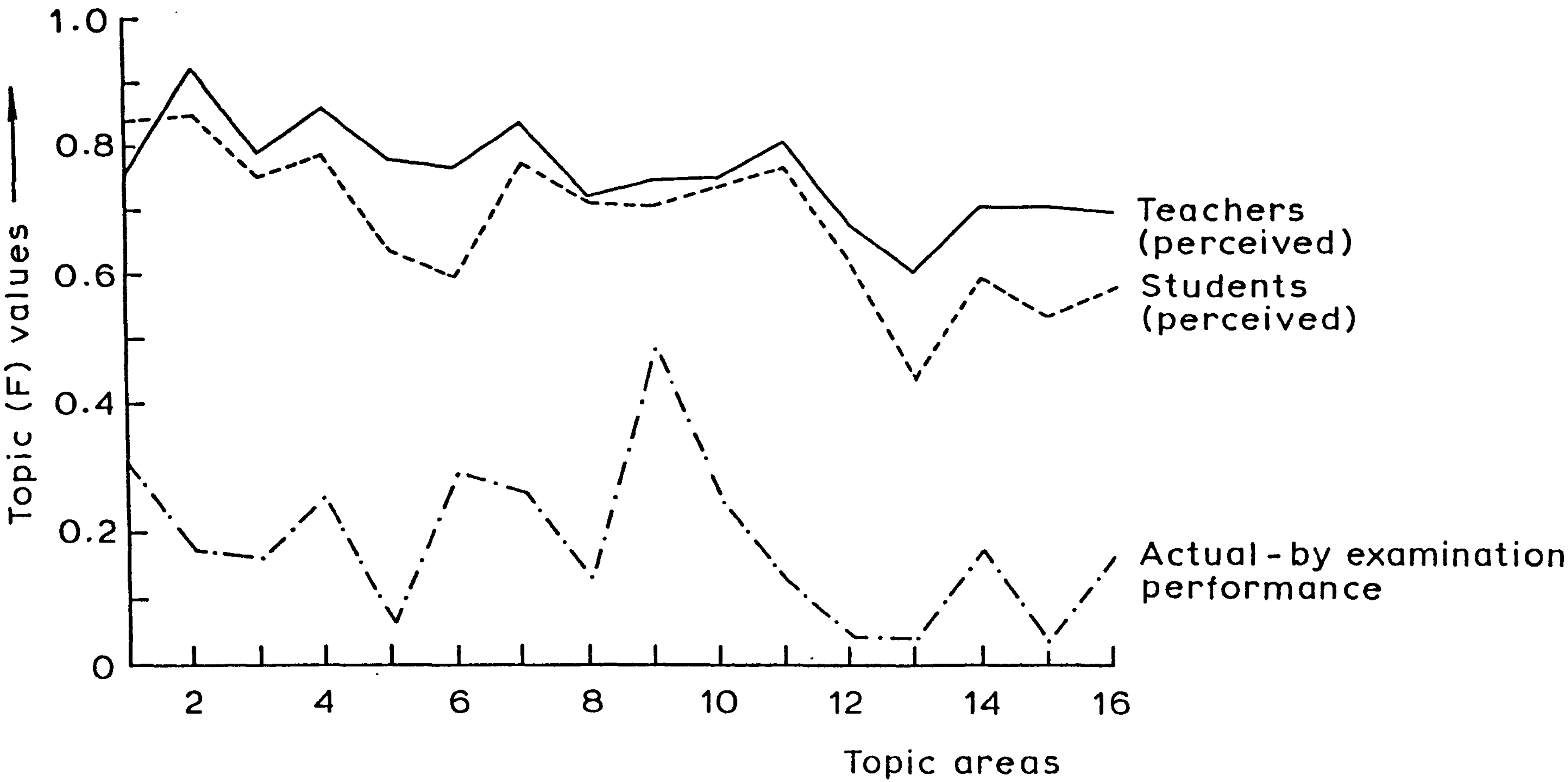
	Perceived by TEACHERS	Actual Exam. results
Perceived by Students	.8035	.4702
Perceived by Teachers		.3329

5.71 Discussion.

As it can be seen from Table 5.24 with regard to the perception of difficulty of various topics, and subtopics of the curriculum, the agreement between teachers' and students' correlates to the magnitude of r = 0.8035. The overall correlation with actual performance was 0.4707 (students), and 0.3329 (teachers). These correlations were all statistically significant beyond the 0.05 level of significance.

Finally, Fig. 5.05 shows the relationship between students' and teachers' perceived topic ease or difficulty and the students' actual topic performance. Some topics which were thought to be easy, for example topic No.2, (Cell structure and function), and No.7 (Excretion and Osmoregulation in animals) which were ranked easiest by both teachers' and students' views, did not give rise to high performance in the examination; while topics which were thought to be difficult, particularly by students, for example topic No.6, (Transport system in plants and animals), which was ranked 13th by students' views, did give rise to high performance in the examination. By the examinations performance, topic No. 6 ranked 3rd, which was closer to teachers' views by which it was ranked 2nd easiest. On the other hand, some topics which were thought to be difficult, for example topic No.5 (Respiration in plants and animals); topic No.13 (Evolution); and topic No.15 (Parasites and diseases of man), did indeed give rise to low performance in the examination. Further discussion of topic difficulty follows in Section 8.83, and Section 10.10.

Fig. 5.05 Relationship between perceived ease or difficulty and actual topic ease or difficulty



CHAPTER SIX

RELIABILITY OF EXAMINATIONS

There are two important qualities that examinations and other measurement instruments should display. These are reliability and validity. The concept of examination reliability is discussed in this chapter, while that of validity is covered in the next chapter.

6.10 Reliability defined.

With respect to a measurement procedure, reliability refers not to 'what it measures' but 'how accurately it measures what it does measure' (Thorndike and Hagen, 1969, p.117). Relevant questions are what is the precision of the score? How accurately will it be reproduced if the individual is measured again? It is an estimation of the consistency among measurements (Cronbach, 1970). That is, the extent to which examination/test score or evaluation results are repeatable (Gronlund, 1971; Nunnally, 1972; Popham, 1975).

There are three things which are notable with regard to reliability: reproduction (repeatability); comparability; and consistency of responses. That being the case, three factors can be said to contribute to examination reliability:

1. The extent to which different forms of the same examination are comparable; that is from year to year.
2. The consistency with which those who take the examination perform; that is throughout the various components of the examination.
3. The consistency with which the examination is marked.

There are many factors which can cause inconsistency in measurement outcomes. Such factors can be grouped under two main categories: Firstly, those operating within the individuals being measured (internal factors); Secondly, those operating externally on the individuals being measured (external factors).

6.20 Factors affecting Reliability of examinations:

Internal factors.

Basically these factors, operating within a person cause a change in the person. The change can be due to one or a combination of any of the following: physical growth (maturation); fatigue; emotional strains; health; fluctuation of human memory; experience in test taking (Tuckman, 1972). All these and many others can cause test anxiety. These factors could be operating within the examinee, the examiner or both. Some are much more likely to affect reliability than others.

External factors.

External factors can be grouped under four types. Those operating on the physical conditions of the examination room; on the examiner; within the examination itself (nature of the examination) and the methods used to estimate the reliability (See the Work of Backhouse, 1972 , which has been reviewed in Section 6.30). There is no guarantee that all the physical conditions in an examination situation can be kept constant in a particular sitting and in a repeat of the examination.

If an examination showed a high reliability in an experimental situation, there is no guarantee that in the actual (real) situation, the experimental conditions can be maintained. And even if all the examination conditions were controlled, still the internal factors, pointed out in the previous paragraph might be in operation.

Inconsistency during scoring is not very common in the scoring of objective items because the correct responses for such items are predetermined and are objective. Inconsistency, however, has been found to be common in the scoring of subjective items especially those of the essay type. Hartog and Rhodes, (1935, 1936) described two types of scoring inconsistency with regard to essay questions: 'intermarker reliability' - between markers marking the same subjective item; and 'inter-occasion reliability' - between same marker marking the same subjective item on two different occasions. Willmott and Nuttall call the inter-occasion reliability 'intra-examiner reliability'. (Willmott and Nuttall, 1972). Harris (1960) refers to the intermarker reliability as 'reader reliability' and defines it as 'the extent to which different individuals can agree on the score which should be assigned a particular essay question'. Gronlund call this 'objective reliability' (Gronlund, 1971) and defines it as 'the degree to which equally competent scorers obtain the same results'.

With regard to the nature of the examination, there are two main features which can lead to inconsistency in the scores: length of the examination; and inherent difficulty of the examination.

The longer the examination, in terms of number of items, the higher the reliability (Ebel, 1970). This is due to the fact that a longer examination provides a more adequate sample of the behaviour being measured and the scores are less apt to be distorted by chance factors such as guessing.

The difficult level of an examination will also affect reliability because, examinations which are too difficult for the group members taking it will tend to provide low achievement scores. Examinations that are too easy also reduce reliability because, as with examinations that are too difficult, a restricted spread of score occurs.

6.30 Reliability of Examinations - a review.

Extensive study has been made of the reliability of examinations, either of reliability itself, for example, Hartog and Rhodes (1935, 1936), or in wider studies incorporating different aspects of the functioning of examinations (Willmott and Nuttall, 1972). Most of the studies on reliability reviewed in the next pages are on British examinations because the British examinations have a number of similar features to the Tanzanian examinations.

One of the earliest examples of studies into the reliability of examination scores are the investigations into the intermarker, and inter-occasion reliability of examination scores carried out by Hartog and Rhodes, (1935, 1936) who arranged for a multiple marking of examination scripts by experienced examiners under examination conditions. Their investigation was on the marking of school certificate history scripts.

To investigate the intermarker reliability, fifteen scripts which had been awarded the same mark by an examination board were selected. These were re-marked by fifteen selected examiners. The examiners re-assigned awards - credit, pass, fail - to the scripts. Hartog and Rhodes observed that the awards given to the scripts differed greatly from the original board awards. Also the awards by each examiner differed.

In another study, to investigate the inter-occasion reliability, the same scripts were given to the same examiners who had not kept a record of their previous assessment. They observed that there were discrepancies in the credit - pass - fail allocation of awards between the same examiners marking the same scripts on different occasions. In both the intermarker and inter-occasion studies there were disagreements as to where the pass - fail, and credit - pass lines were; and differences in relative merit.

Since that pioneering work, many studies into the reliability of examination scores have been carried out, not least by examination boards themselves although the results of this work are not often published but tend to be used as part of the constant process of refining and improving the technical efficiency of their assessment procedures. The examples which follow have been picked to illustrate cases where biology examinations have either been the centre of the investigation or have been included as part of a more general one.

Backhouse (1972) compared methods of estimating examination mark reliability in situations where a choice of questions is allowed. He covered both the ordinary and advanced level examinations of the G.C.E. in England and Wales, in different subjects. He used seven different formulae for the estimation of each examination mark reliability. In an A-level Biology Paper for four combinations of questions which were most frequently answered by the candidates, the following reliability values, using Cronbach alpha, were observed: 0.82 (n = 114); 0.78 (n = 65); 0.80 (n = 61); and 0.87 (n = 53).

For O-level Biology, he observed different values for each of the seven methods used on each paper and part of paper. The values ranged from 0.739 to 0.830 and 0.751 to 0.829 for paper I and paper II respectively for the different methods for one sample (n = 174). Using a different sample (n = 175), he reported values ranging between 0.717 and 0.848 and from 0.714 to 0.824 for Paper I and II respectively. In discussing his results, Backhouse concluded that, the range of values observed was a 'creditable achievement' on the part of the examination boards, and that 'with such selected populations it would, indeed, be surprising to obtain a reliability greater than 0.9' (ibid.p.113). He also points out that practical papers tended to yield lower reliabilities than theory examinations.

Willmott and Nuttall, (1975) published a study of the reliability of 83 British CSE examinations held in 1969 and 1970; and 29 GCE O-level examinations held in 1970 and 1971. A total of ten examination subjects were studied: maths, English, geography, biology, woodwork, commerce, French, religious knowledge and physics. For the 1970 CSE examinations in biology, the values observed ranged from 0.78 to 0.94 with a median value of 0.88. 'All the median values equal or exceed 0.85 and the results may therefore be considered to be most encouraging, though not incapable of improvement' (Willmott and Nuttall, 1975, p.23).

Willmott and Hall (1975), while examining the effects of question choice in O-level examinations, studied a number of different boards' examinations. As part of the study, they analysed reliabilities of the examinations studied. For six cases of GCE O-level biology examinations studied, they reported reliability values ranging between 0.75 and 0.83 with a median value of 0.79. (ibid. p.126).

The Schools Council in England, published a report by Duckworth and Hoste (1976) on 'Question banking: an approach through biology'. They studied the construction of pre-tests and the analysis of items, and the reliability and validity of a broad spectrum of question types when used to test the understanding of biology gained at 16-year-olds' public examinations.

The total number of pre-tests studied was 24. The reliabilities of the pre-tests, established by the Kuder-Richardson formula 20 (KR20), observed by the researchers ranged from 0.65 to 0.77. The outcome of the analysis of the items was that a large proportion of the multiple-completion and assertion/reason items were judged unsuitable for incorporation into any working bank of biology questions for 16-year-old pupils. On the other hand, short-answer open-ended questions and straight forward multiple-choice questions seemed, on the whole, more suitable except that most of them were judged to test the knowledge objective only.

In the same study, the effects of using three different marking schemes to mark the same essay question was examined. One of the schemes was highly structured. The results of this study showed that the mean scores awards by all the markers were below the mid-point of the intended mark range, whichever scheme was used. The highly structured mark scheme gave the lowest weighted mean of the means. An intermarker correlation, for 26 marker pairs, using the different schemes range from 0.69 to 0.91. Again the most highly structured scheme did not give the most reliable intermarker correlation as had been expected. The scheme had restricted the markers to a narrow range of marks without any compensating advantage, such as increasing the intermarker reliability of the markers using it.

In an evaluation of a Certificate of Secondary Education Examination in Biology, Hoste, (1977) studied one of the CSE Board's examination for 1972. The examination studied was a typical examination with written and practical papers. The items on the papers were also of different types, multiple-choice, matching, true-false, short-answer, and essay type items.

Hoste studied the reliabilities of the whole examination, and of the different papers and components of the different papers. The reliability values observed, using Cronbach's coefficient alpha, range from 0.568 to 0.891. He concluded that the estimates of reliability were 'within the range found in other CSE examinations' (ibid, pp.125-130).
Furthermore, he pointed out that the components of the examination containing constructed response items were consistently more reliable than the component of the examination consisting entirely of multiple-choice questions.

As can be seen from the studies reviewed above, the reliabilities reported are of different orders. In almost all cases an element of satisfaction, with regard to the results, is expressed.

The study of the reliability of the Tanzanian O-level biology examination reported from Section 6.50 of this Chapter is based on an analysis of an examination which consisted of the following types of items - essay, multiple-choice, matching, true-false, and short-answer free-response items included in the Theory and Practical papers (See Appendix IIIa and IIIb).

The marks which constitute the data are marks assigned to responses to the items at a single typical marking session which was part of an operational examination, each script having been marked by a single examiner. The possible methods for estimating the reliability of examinations are discussed in the next Section and the reasons for selecting the method used in this study are given.

6.40 Methods of Estimating Reliability.

There are several methods of determining reliability which are discussed in most texts on educational and psychological measurement. These are: test-retest method; equivalent form method (parallel test form); internal consistency method; (split-half approach, and Kuder-Richardson approach). These methods are briefly discussed below.

Test-retest method.

By this approach, the same examination/test or instrument is given twice to the same individuals, with a time interval between the two administrations of the test. This method is a measure of stability of the examination, test or any other evaluation instrument. If the examination is stable, high correlation coefficients should be observed between the scores of the first administration (pre-test) and scores of the second administration (post-test).

The most preferable time interval should be determined by the use to be made of the results. For example, as part of this study, the reliability of the Student Opinion Questionnaire was estimated through a test-retest approach with a time interval of 25 days. For details see Chapter 4 Section 4.60.

Equivalent forms (Parallel test forms).

This involves the use of two different but equivalent forms of an examination. The two forms are administered to the same group of students in close succession and the resulting scores are correlated. If a time interval is allowed between the administration of the two forms, they can measure both stability and equivalence. In this respect they will be 'test-retest equivalent forms'.

Internal consistency method.

By this method, reliability of the examination is calculated from scores of a single examination and a single administration to a group of candidates. There are several varieties of internal consistency methods of estimating reliability of the scores obtained on the examination. Most of these varieties consider the relationship between individual questions as a means of estimating the reliability of the total scores obtained on the examination. Most common of these are: the split-half approach; and the Kuder-Richardson approach. Most other approaches found in educational measurement texts are a modification of one or both of the above (for exampleCronbach, 1970; Willmott and Nuttall, 1972). Cronbach alpha is used to study the reliability of the Tanzanian O-level Biology examination in the later part of this Chapter, (See Section 6.51).

The basic principle behind the split-half approach is to split a single examination, administered at a single sitting into two halves by splitting the items into two lots. The two half tests can either be marked as one examination or separately. Several methods of splitting the examination into the two halves can be employed: Using odd and even numbered items; the first and last halves; or random selection of questions to form two equal groups. The most commonly used approach is the selection of odd and even items. No matter which method of division is chosen, the basic property of the split-half method of estimating reliability will hold.

The Kuder-Richardson approach has two basic formulae for the estimation of reliability: the K-R20; and the K-R21. K-R20 is based on the proportion of candidates gaining credit on each item and the standard deviation of the total scores. The resulting coefficient is equal to the average of all possible split-half coefficients for the group examined. The essential assumption in the procedure is that the items within one form of an examination have as much in common with one another as do the items in that one form with the corresponding items in a parallel or equivalent form. That is, it is assumed that the items are homogeneous, every item measuring the same general factors or whatever is being measured. Clearly this is the major shortcoming for the use of K-R20 for the estimation of the reliability of an examination. Achievement tests and examinations measure different types of learning outcomes. Therefore the items on an examination are unlikely to be homogenous. If K-R20 is used to estimate the reliability of examinations, it should be born in mind that there will be an underestimation of the reliability.

K-R21 is much more simplified and can only be applied to the results of objective tests which have been scored on the bases of the number of correct answers: the multiple choice; true-false; matching and the single word or phrase answer items.

The method of internal consistency described above should not be used with speeded tests. It should also be noted that the internal consistency procedures do not indicate the constancy of candidate responses from day to day as the estimate is based on a single examination administered once to a group.

6.50 ESTIMATION OF THE RELIABILITY OF THE TANZANIAN O-LEVEL BIOLOGY EXAMINATION OF 1983.

6.51 Data analysis and results.

As already pointed out, one of the purposes of this study was to obtain information about the reliability of the Tanzanian O-level Biology examination. To obtain this information, each question, and each component of the examination was analysed in detail. The scripts included in the analysis include all the 519 scripts used in the previous chapter, plus an additional 296 scripts. The details of the sampling procedures have already been discussed in Section 4.51.

All the questions on the Theory paper, except for the three essay questions, and all the questions on the Practical paper were compulsory.

Cronbach coefficient alpha, which is equivalent to KR-20, (Kuder-Richardson 20), was used to estimate the reliability. The shortcoming mentioned earlier was minimised by performing the analysis on separate sections of the papers within which the items could be thought to be reasonably homogeneous. The strength of this assumption is tested later in this work (see Section 8.71 and Section 8.73), when the demands of the various components on the candidates are examined more closely in assessing the validity of the examinations. The formula for Cronbach's coefficient alpha given in 'Subprogramme Reliability - The Model = Alpha Specification' of SPSS is:

$$\text{coefficient alpha } (\alpha) = \frac{k}{k-1}\left(1 - \frac{\sum_{i=1}^{k} S_i^2}{S_T^2}\right)$$

where k = number of items involved.

S_i^2 = the variance of the item i.

S_T^2 = the variance of the sum over the k items.
(Nie and Hull, 1981)

A further analysis, to estimate the standardized coefficient alpha, was carried out using the formula:

$$\text{standardized coefficient alpha } (\alpha) = \frac{K\bar{r}}{1 + (K - 1)r}$$

where r = the average correlation between items.

In essence, standardized coefficient alpha is calculated when the scores on each item are standardized by dividing them by the standard deviation of the item. All the computations were done through 'Subprogramme Reliability - The Model = Alpha Specifications' of the SPSS (Nie et. al, 1981).

Information relating to the Theory examination as a whole, and its components is given in Table 6.01, that relating to the Practical examination and its components is given in Table 6.02. The tables show maximum marks obtainable for each part of the examination as well as for the whole, covering the mean, standard deviations, coefficient alpha, and standardized coefficient alpha.

6.52 Discussion.

As can be seen from both Tables 6.01 and Table 6.02, in each of the parts, except in the true-false component of the Theory paper, the mean mark is much lower than 50 per cent of the maximum possible mark obtainable. Getting the mean as near to 50 per cent as possible maximizes the chances of achieving a good spread of scores and for most situations is considered the ideal position (Willmott and Nuttall, 1972).

The sizes of the standard deviations, ranging from 0.79 to 15.11 are more important in that, if the marks are well spread, the chances of achieving satisfactory reliability are greatly increased.

It is generally accepted that a standard deviation which is one-seventh or more of the maximum possible mark, represents a very 'satisfactory spread of scores' (Nuttall, 1972). If the standard deviation is a smaller fraction than one-seventh, then there is considerable room for improvement.

Looking at Table 6.01 and Table 6.02, the standard deviations of the Theory examination as a whole, the essay section, the free-response section, the Practical examination as a whole, and the observational investigation section are all smaller than one-seventh of the maximum possible mark. This could mean that there is room for improvement in these components and the examination as a whole. On the other hand, the standard deviations of the multiple-choice, matching, true-false (in the Theory paper), and experimental section (in the Practical paper), are all greater than one-seventh of the maximum possible mark, hence satisfactory. However, the standard deviation for the true-false section is only slightly higher, (.792), than one-seventh of the maximum possible mark, (0.714).

With regard to reliability, Table 6.01, and Table 6.02 show that the reliabilities of the two examination papers, (the Theory and the Practical) are of different order. The reliabilities of the Theory examination components are lower than those of the Practical examination components with values

Table 6.01

The estimates of reliabilities of the Tanzanian 1983 O-Level Biology Examination.

The Theory Examination.

Paper Component	Maximum Possible Mark	Mean	Standard Deviation	Reliability	
				Alpha	Standardized Alpha
Essay	20	3.68	2.523	.474	.289
Multiple Choice	5	2.48	0.932	.661	.631
Matching	5	1.90	1.122	.710	.752
True-false	5	2.87	0.792	.594	.486
Free-response	65	10.14	5.601	.711	.887
Entire Theory Paper	100	21.07	8.510	.732	.836

Table 6.02

The Practical Examination.

Paper Component	Maximum Possible Score	Mean	Standard Deviation	Reliability	
				Alpha	Standardized Alpha
Experimental Investigation	53	9.68	12.160	.715	.952
Observational Investigation	47	14.29	5.691	.709	.809
Entire Practical Paper	100	23.97	15.108	.829	.846

ranging from 0.474 to 0.711; while those of the Practical examination range from 0.709 to 0.715. The reliabilities of the entire Practical paper are slightly higher than those of the Theory paper: that is (0.829 and 0.846); and (0.732 and 0.836) respectively. Although the true-false section of the Theory paper showed the highest (F) values, (ref. Table 5.13), relative to the other components and the entire Theory examination, it has one of the lowest reliability estimates; the lowest being the essay section, presumably because of the relatively small number of items involved and also their low standard deviations.

The reliabilities, and all the other indices given in table 6.01 and Table 6.02 are related to marks and not to grades awarded to the candidates. Ebel (1965), in discussing the grading process, points out that there is always a loss of information, leading to a loss of accuracy in the process of translating marks to grades. The grades awarded to candidates using the marks obtained on the biology examination is beyond the scope of this study. Therefore, it is given no further discussion here. However, since the effectiveness of the grading process rests, to an extent on reliability, the low reliability values quoted above, for some of the sections, must give some cause for concern.

CHAPTER SEVEN

VALIDITY OF EXAMINATIONS: THE CONCEPT OF CONTENT VALIDITY.

7.10 Introduction.

Validity, probably the most fundamental criterion of the worth of an examination, is a concept which has been asigned a number of interpretations by different authors and researchers.

The concept is sometimes talked of in terms of effectiveness, for example, Brereton defines validity as a concept used to 'express the effectiveness with which an examination or test measures that which it is intended to measure, or the effectiveness with which it accomplishes the purposes it is intended to accomplish' (Brereton, 1944, p.63).

Cronbach, on the other hand, defines the concept of examination validity in terms of its predictive power. He agrees with Catell and Warburton (1967), and defines examination validity as 'the ability of a test to predict something other than itself' (Cronbach, 1970, p.121).

Validity, also, refers to 'the extent to which the results of an examination procedure serve the particular uses for which they are intended (Gronlund, 1971, p.75), while Hoste defines validity in terms of the congruence between what should be tested and what is tested (Hoste, 1977, p.131).

With the element of evaluating results in mind, the concept of validity, therefore, pertains to the results of an exam/test/or evaluation instrument, and not to the instrument itself. We should, therefore, speak of the validity of examination results or validity of results of an evaluation instrument and not the validity of an examination or evaluation instrument per se (Gronlund, 1971).

Validity is a matter of degree. It does not exist on an all-or-none basis. Results should not be thought of as being valid or invalid, but in terms of categories that specify degree, such as high, average, and low validity; and most importantly it should be remembered that validity is always specific to some particular specified use. When appraising or describing validity therefore, it is necessary to consider that, the answer to the question, 'Is this test valid?', should be answered by a return question, 'Is it valid for what?' (Connaughton, 1969, p.164; Guilford and Fruchter, 1978, p.435).

Examination results then are never just valid, they have a different degree of validity for each particular use to which they are put. It does not mean that if an examination result is valid for making one decision, it is a guarantee that it will be valid in making other decisions. The uses which can be made of examination results have already been discussed in Chapter 3 and the area is discussed later in the light of the analysis undertaken.

7.20 <u>Types of Validity</u>.

The relativity of the concept of validity discussed in the previous section, is further complicated by the existence of several varieties of validity. Most educational measurement texts distinguishes three basic varieties of validity:

1. Criterion-orientated validity (also referred to as predictive validity or concurrent validity).
2. Content validity.
3. Construct validity.

The differentiation of the validity concept into the above categories originates from a recognition of the various uses made of tests and psychological aids.

Criterion-oriented validity tries to explain how measures of some valued performance, (criterion), relates to the test scores to predict or estimate the criterion. For example, the use of intellectual aptitude tests to predict how well students will succeed in further education.

Content validity explains the extent to which an examination measures a representative sample of the subject matter content and/or the abilities under consideration.

Construct validity tries to explain how scores on a test can be explained in terms of psychological constructs. A construct can be defined as a factor or a 'characteristic that is believed to account for some aspect of human behaviour, such as intelligence, sex, age, attitude' (Turney and Robb, 1971, p.155). Cronbach and Meehl (1955), define a construct as 'some postulated attribute assumed to be reflected in test performance. In test validation, the attribute about which we make statements in interpreting a test is a construct'. Each construct has an underlying theory which can be brought to bear in describing and predicting a person's behaviour.

Several other varieties of validity have been described. Ebel (1972), lists 10 different varieties of validity which have been described by different writers: concurrent; construct; content; curricular; empirical; face; factorial; intrinsic; predictive and validity by definition. Ebel's classification of the 10 varieties into two major categories: those concerned with primary or direct validity; and those concerned with secondary or derived validity is useful.

A test has direct primary validity to the extent that the tasks included in it represent the kind of tasks that provide an operational definition of the achievement or the trait in question. It has derived secondary validity to the extent that the scores it yields correlate with criterion scores which possess direct primary validity (Ebel, 1972, p.438). Ebel's classification of the 10 varieties into the two categories of primary and derived validity is summarized in table 7.00.

Table 7.00

Categories and varieties of validity

DIRECT (PRIMARY) CATEGORY	DERIVED (SECONDARY) CATEGORY
Content by definition	Empirical validity
Curricular validity	Concurrent validity
Validity by definition	Predictive validity
Intrinsic validity	Factorial validity
Face validity	Construct validity

The above varieties are not all distinctly different from each other. Even the distinction between the two major categories is not sharp in all cases.

Whether direct or derived validity, two factors dominate the process of validation: Judgement and evidence. Rational and professional judgement, and empirical and statistical evidence cannot totally be separate elements in validation, they are interconnected. As will be seen later in this chapter and the next, both judgement and empirical evidence are the basis for validation of the Tanzanian O-level biology curriculum and Examination scores. Two varieties of validity, content and construct have been studied in detail.

7.30 Studies of Examination Validity

There are relatively fewer research works which have reported on the validity of examinations as a whole, than on the reliability of examinations. Most of the studies which have been reported have mostly concentrated on content validity. As pointed out in the previous section, content validity of an examination explains the extent to which an examination measures a representative sample of the subject matter content and/or the abilities under consideration.

Most studies which have been carried on the content validity of examinations have been approached from the point of view of coverage of abilities rather than subject matter coverage. For example Crossland and Amos (1965), quantified the emphasis placed on four outcomes of biology teaching by studying the content of the O-level examination papers of nine GCE examination Boards. The papers were O-level biology, general science; chemistry and physics for the period 1948 to 1964.

The examination questions were categorized according to four learning outcomes: the acquisition of facts; the interpretation of facts and the drawing of conclusions from experiments; the application of scientific principles to new situations; and the designing and planning of experiments.

In all four subjects studied, Crossland and Amos reported that the examinations placed the greatest emphasis on testing the acquisition of facts and least emphasis on the interpretation and drawing of conclusions. For the biology examinations, they reported the emphasis placed on acquisition of facts to be as high as 90% of the total coverage.

Upon a scrutiny of the syllabuses of the subjects issued by the nine GCE Examining Boards they further pointed out that, apart from one Board, references to scientific method were limited in the syllabuses. Therefore, they concluded that 'scientific method is virtually ignored in O-level science courses' (Crossland and Amos, 1965, p.36).

One of the systems of classifying educational objectives which has been widely used to determine the content validity of examinations is Bloom's Taxonomy of Educational Objectives (Bloom, 1956). Bloom's Taxonomy lists six hierarchical categories of conceptual development. The six categories, from lowest to highest level, are: Knowledge; Comprehension; Application; Analysis; Synthesis and Evaluation. A short description of each of the six categories, including examples of statements of instructional objectives is given in Appendix IVe. Appendix IVe was developed specifically for use in a later section of this study.

Lewis (1967) studied multiple-choice tests of ability in three subjects: physics, chemistry and biology. The multiple-choice tests were constructed to assess the attainment of three learning outcomes: basic knowledge; comprehension and application; and evaluation. The tests were administered to a group of boys who had completed the ordinary level science syllabus course of the Joint Matriculation Board (JMB).

The results were factor analysed and the analysis suggested that, although 'evaluation' was an important aspect of the abilities tested, it was not developed through O-level courses since candidates performed poorly in this category. Lewis's findings agree with those of Crossland and Amos, (1965), in that the O-level school science courses were concerned 'first and foremost with preparation for the O-level examinations, examinations in which it has been noted to place emphasis on the lower level aspects of Bloom's Taxonomy' (Lewis, 1967).

Willmott and Hall (1975), as part of a larger study, analysed the content validity of some GCE and CSE O-level examinations through a qualitative approach. The aim of the study was to see if the overall content of questions, in the form of a specification of educational objectives, was related to other aspects of the examinations, such as examination facility, reliability and comparability.

Willmott and Hall condensed Bloom's Taxonomy of six hierarchical categories into four categories: Knowledge; Comprehension; Application and Analysis/Evaluation. A total of seven examinations in three subjects, geography, biology and physics, from four Examining Boards, were studied. Subject panel members were instructed to classify each part of each question according to the various levels of ability as defined by the modified Bloom's Taxonomy.

As regards the classification of the questions, they reported that 'it was evident from all the examinations analysed that the degree to which members agreed in their classification varied considerably from question to question, although there was a far greater consensus of opinion on questions with a high content of knowledge than on questions where a wider range of ability levels were tested' (Willmott and Hall, 1975, p.110).

The striking feature of the analysis by Willmott and Hall was the consistency of the total classification of questions across the seven examinations analysed. As with Crossland and Amos (1965) and Lewis (1967), Willmott and Hall reported that the contribution of knowledge and comprehension to the overall examination was high, ranging between 54 and 66 per cent, and from 22 to 36 per cent respectively. They concluded that recall played a predominent role in determining the performance of candidates in O-level examinations.

Furthermore, they reported that the taxonomic headings did not appear to be related to the facility values for the questions, but, they observed that the more able candidates selected the easier questions, which in general were questions identified as having a larger content of knowledge.

Fairbrother (1975) was interested in finding out whether there was any agreement among teachers about the abilities which were being tested by the 1970 and 1971 multiple-choice papers of the Nuffield Advanced Physics examination.

The teachers' task was to classify the items according to whether they felt the items tested knowledge comprehension, application and analysis/ evaluation.

Fairbrother observed that there was no coefficient of agreement reached and it was finally decided that 'each teacher is a judge in his own right and we have no reason to accept the opinion of one rather than of another' (Fairbrother, 1975, p.203).

However, the general opinion pointed more to the items placing emphasis on the lower level abilities (knowledge and comprehension) (72.9%) and less emphasis on application and analysis, (27.1%), for the 1970 examinations. The same trend was observed for the 1971 examinations.

Another aspect of examination validity which researchers have started to consider is construct validity. Construct validity tries to explain how scores on a test can be explained in terms of traits other than those specified directly in the examination questions. The review of studies which have looked into construct validity of biology examinations, and the analysis of the construct validity of the Tanzanian Biology examinations of 1983 currently under study is covered in the next chapter.

As pointed out in an earlier section of this chapter, there are many different ways of conceptualising validity. With so many different usages and interpretations, it is clear that no researcher can study all aspects of validity at once.. For many educational achievement tests, the most important type of validation is direct validation, and this may sometimes be the only possible alternative (Ebel, 1972). This is the true of the present study, and the next sections of this chapter will consider the content validity of the Tanzanian 1983 O-level biology examinations, covering both subject matter and abilities tested.

7.40 The Content Validity of the 1983 Tanzanian Biology Examinations.

In order to carry out any type of direct validation, it is necessary to define the achievement or trait to be measured. In this study, the traits to be measured have been defined in two ways, firstly, in terms of the abilities intended to be developed during the four year course. Thirteen (13) objectives have been identified, by summarizing those listed in the Teachers' Guide (1978). The list of the 13 objectives is shown in table 7.01.

Secondly, the instructional content areas to be covered have been identified as sixteen (16) main topic areas drawn from the O-level Biology syllabus (1976), and the Teachers' Guide.

The analysis of the attainment of the objectives was preceded by a study of the importance attached to the objectives in the O-level biology course and in the examinations by O-level biology teachers concerned with preparing students for the examination, and by biology panel members, concerned with preparing the curriculum.

Table 7.01.

The list of course objectives prepared by summarizing those listed in the Teachers' Guide (1978). (See Appendix II).

	Objective
1.	Ability to:- Show knowledge and understanding of biological facts, concepts, principles and theories.
2.	Comprehend biological texts.
3.	Apply biological principles to real life experiences.
4.	Plan experiments.
5.	Follow instructions in carrying out experimental and other biological procedures.
6.	Make and record observations accurately.
7.	Interpret data to draw conclusions and make inferences.
8.	Draw biological diagrams accurately.
9.	Analyse and solve problems.
10.	Formulate hypotheses from numerical data, graphs, tables and other evidences.
11.	Explain and discuss.
12.	Write accurate reports of experiments and other biological procedures.
13.	Use manual skills to set up and operate laboratory apparatus.

7.41 The rating and ranking of curriculum and examination objectives.

Many studies have reported on the identification, and/or rating of objectives in biology. They are of two kinds. First, there are those which have examined the general objectives of O-level and A-level biology curricula (Amos, 1970; Duckworth and Hoste, 1976; Jungwirth, 1971; Carrick, 1983); and secondly, those which have looked specifically into the objectives of practical biology (Abouseif and Lee, 1965; Head, 1966; Lister and Kelly, 1969; Tamir and Glassman, 1971; Thompson, 1975; Lynch and Ndyetabura, 1982; Beatty and Woolnough, 1982).

The studies covering the general objectives of teaching biology, are reviewed below, while those covering the objectives of practical biology are reviewed in the next chapter where practical biology is studied in relation to the concept of construct validity.

Amos (1970) reported the use of a postal questionnaire to obtain the opinion of biology teachers, in the northern half of England on the objectives of biology teaching up to O-level GCE.

The questionnaire contained a list of nineteen objectives. Amos was interested in finding out teachers' opinions as a result of rating the objectives on a five-point-scale according to the importance they felt they had: in their actual teaching; in ideal circumstances of the teacher's own choice; and in the GCE O-level examination.

In teaching, Amos reported that the objectives concerned with acquisition of knowledge were the prime concern in their actual teaching, while objectives concerned with scientific method were rated lower. In the ideal situation, the order was reversed, so that scientific method was considered more important than acquisition of facts. However, the order of priority for the GCE examination resembled markedly that in teaching. This suggests that teachers will often stress the same things which are stressed and demanded by the examinations, otherwise they would be unfair to their pupils if they did not adopt this strategy.

In a similar study, Duckworth and Hoste, (1976), circulated a list of 15 subject areas and 14 curriculum objectives to experienced item writers and teachers to assess the importance of the subject areas and curriculum objectives in teaching and examinations. The subject areas and curriculum objectives were identified by means of a scrutiny of CSE and GCE examination boards' syllabuses, from reviews of research, and through discussions with teachers.

Duckworth and Hoste observed only small differences between the opinions of item writers on one hand, and teachers on the other hand in both subject areas and the objectives.

In another study, Jungwirth (1971) studied teachers' and pupils' priorities of objectives in the Israel version of BSCS. A list of 14 statements of educational objectives of biology teaching was prepared by adapting statements from various BSCS sources. The teachers were asked to rank the objectives in order of their relative importance; and re-rank the objectives in order of their feasibility of being achieved by pupils studying BSCS biology in the 9th and 10th grades.

The same list of objectives was given to 11th grade pupils who had been taught by the above teachers in the 9th and 10th grades. The pupils were asked to rank the objectives in order of their relative importance as seen by themselves; and re-rank the objectives in order of their importance as they would expect their biology-teacher to rank them.

Jungwirth observed little concensus of opinion, among the teachers, most objectives being assigned high priorities by some and low priority by others. He also observed that teachers priorities were not identical with their expectations, although there was a significant correlation between the two (0.75).

With regard to pupils' opinions, he observed that pupils' ranking and ranking of the 'teachers image' were very similar, with an overall correlation of 0.92. The pupils' opinions identified closely with their assumptions of the biology teachers' opinions.

7.42 <u>The rating of the objectives of the Tanzanian O-level Biology course</u>.

The list of 13 biology course objectives prepared by summarizing those listed in the Teachers' Guide (1978), (See table 7.01), was presented to fourth form biology teachers concerned with teaching and preparing students for the examination, and biology panel members, concerned with developing, monitoring and changing the curriculum. Forty nine (49) out of the eighty five (85) questionnaires circulated to teachers were returned duly completed. At the time of the study, the Biology Panel consisted of 10 members. Eight (8) out of the ten (10) members returned completed questionnaires.

In the determination of the importance of the objectives, the biology teachers and panel members were required to rate and rank the list of the 13 objectives firstly, according to the importance in an O-level biology course using the key: of no importance (1); of some importance (2); very important (3).

Secondly, the teachers and panel members were asked to rank order the objectives from 1 to 13 on the basis of the importance and emphasis they felt should be given in an O-level examination. That is, according to the abilities that an O-level biology examination would test.

It was hoped that an analysis of the rating procedures would indicate the relative satisfaction felt in the representation of the objectives in the course, and a comparison of the rating and ranking of the objectives would provide interesting indications of the perceptions held by the respondents of the validity of the examinations. If the rank order of the objectives when rated as course objectives was identical to the order in which the objectives were ranked for examination purposes the examination would appear to enjoy a high validity in the eyes of the teachers. The rating and ranking of the objectives by the teachers and panel members is discussed in the next section.

Table 7.02

Teachers' rating of the course objectives according to importance in O-level biology course.

	Objective	Mean rating	range of rating (1-3)	rank order from mean rating
	Ability to:			
1.	Show knowledge and understanding of biological facts, concepts, principles and theories.	2.878	2-3	1
2.	Comprehend biological texts.	2.327	1-3	12
3.	Apply biological principles to real life experiences.	2.878	1-3	1
4.	Plan experiments.	2.367	1-3	10
5.	Follow instructions in carrying out experimental and other biological procedures.	2.694	1-3	5
6.	Make and record observations accurately	2.755	2-3	3
7.	Interpret data to draw conclusions and make inferences.	2.714	2-3	4
8.	Draw biological diagrams accurately	2.612	2-3	8
9.	Analyse and solve problems.	2.673	1-3	7
10.	Formulate hypotheses from numerical data, graphs, tables and other evidences.	2.347	1-3	11
11.	Explain and discuss.	2.551	1-3	9
12.	Write accurate reports of experiments and other biological procedures.	2.694	1-3	5
13.	Use manual skills to set up and operate laboratory apparatus.	2.306	1-3	13

Table 7.03

Teachers' ranking of the course objectives according to ability that an O-level biology examination should test.

	Objective	Mean ranking	Range of ranking (1-13)	Rank order from mean ranking
	Ability to:			
1.	Show knowledge and understanding of biological facts, concepts, principles and theories.	1.711	1-11	1
2.	Comprehend biological texts.	8.500	1-13	11
3.	Apply biological principles to real life experiences.	2.500	1-12	2
4.	Plan experiments.	8.304	2-13	9
5.	Follow instructions in carrying out experimental and other biological procedures.	5.333	2-12	4
6.	Make and record observations accurately	6.000	2-11	6
7.	Interpret data to draw conclusions and make inferences.	5.667	1-13	5
8.	Draw biological diagrams accurately	8.756	2-13	13
9.	Analyse and solve problems.	6.575	1-13	7
10.	Formulate hypotheses from numerical data, graphs, tables and other evidences.	8.400	3-13	10
11.	Explain and discuss.	8.605	1-13	12
12.	Write accurate reports of experiments and other biological procedures.	8.150	2-12	8
13.	Use manual skills to set up and operate laboratory apparatus.	5.271	1-12	3

Table 7.04

Biology Panel members' rating of the course objectives according to the importance in O-level biology course.

	Objective	Mean rating	range of rating (1-3)	rank order from mean rating
	Ability to:			
1.	Show knowledge and understanding of biological facts, concepts, principles and theories.	2.750	2-3	2
2.	Comprehend biological texts.	2.625	2-3	7
3.	Apply biological principles to real life experiences.	2.750	2-3	2
4.	Plan experiments.	2.375	2-3	12
5.	Follow instructions in carrying out experimental and other biological procedures.	2.750	2-3	2
6.	Make and record observations accurately.	3.000	2-3	1
7.	Interpret data to draw conclusions and make inferences.	2.625	2-3	7
8.	Draw biological diagrams accurately.	2.375	2-3	13
9.	Analyse and solve problems.	2.750	2-3	2
10.	Formulate hypotheses from numerical data, graphs, tables and other evidences	2.500	2-3	10
11.	Explain and discuss.	2.625	2-3	7
12.	Write accurate reports of experiments and other biological procedures.	2.750	2-3	2
13.	Use manual skills to set up and operate laboratory apparatus.	2.500	2-3	11

Table 7.05

Biology Panel members' ranking of the course objectives according to abilities that an O-level biology examination should test.

	Objective	Mean ranking	range of ranking (1-13)	rank order from mean ranking
	Ability to:			
1.	Show knowledge and understanding of biological facts, concepts, principles and theories.	3.857	1-11	2
2.	Comprehend biological texts.	8.500	5-11	10
3.	Apply biological principles to real life experiences.	2.875	1-7	1
4.	Plan experiments.	6.000	3-10	6
5.	Follow instructions in carrying out experimental and other biological procedures.	4.143	1-11	3
6.	Make and record observations accurately.	4.625	3-7	4
7.	Interpret data to draw conclusions and make inferences.	5.625	2-8	5
8.	Draw biological diagrams accurately.	12.143	9-13	13
9.	Analyse and solve problems.	7.000	3-13	8
10.	Formulate hypotheses from numerical data, graphs, tables and other evidences.	6.833	3-13	7
11.	Explain and discuss.	8.333	4-13	9
12.	Write accurate reports of experiments and other biological procedures.	8.750	4-13	11
13.	Use manual skills to set up and operate laboratory apparatus.	10.375	8-13	12

7.43 <u>The results and discussion of the rating and ranking of the objectives.</u>

Tables 7.02 to 7.05 show teachers' and panel members' rating and ranking of the objectives as course and examination objectives respectively. There are differences in the opinions of the teachers. Most of the objectives were assigned all possible ratings and rankings. However, the teachers agreed in the rating and ranking of the following objectives:

1. Ability to show knowledge and understanding of biological facts, concepts, principles and theories, and

9. Ability to analyse and solve problems.

These were rated and ranked as priority number one (1) and number seven (7) respectively.

As far as the panel members were concerned, like the teachers, most of the objectives have been assigned high rating and ranking by some and low rating and ranking by others. The panel members however showed unanimity in rating and ranking objectives:

1. Ability to show knowledge and understanding of biological facts, concepts, principles and theories, and

8. Ability to draw biological diagrams accurately.

These have been rated and ranked as priority number two (2) and number thirteen (13) respectively.

Tables 7.06 to 7.09 show a comparison of the teachers' and panel members' opinions. The perceptions of the relative importance of the objectives which were held by teachers and panel members are not identical. For example, while teachers ranked objective No.13 (ability to use manual skills to set up and operate laboratory apparatus) high (3rd), panel members ranked it relatively low (12th) as a course objective. Also, while teachers ranked ability No.2 (ability to comprehend biological texts) low (12th) as a course objective, panel members ranked it relatively high (7th). And while ability No.8 (ability to draw biological diagrams accurately) was ranked relatively high by teachers (8th), it was ranked least (13th) by panel members as a course objective.

However, the following course objectives were assigned the highest five (5) ranks by both teachers and panel members. Ability to:-

1. Show knowledge and understanding of biological facts, concepts, principles and theories.
3. Apply biological principles to real life experiences.
5. Follow instructions in carrying out experimental and other biological procedures.
6. Make and record observations accurately.
7. Interpret data to draw conclusions and make inferences.
9. Analyse and solve problems.

With the exceptions of objective No.1 the above listed abilities are all concerned with scientific method. However, objective No.4 (ability to plan experiments) and ability No.10 (ability to formulate hypotheses from numerical data, graphs, tables and other evidences), also concerned with scientific method, were rated relatively low by both teachers and panel members.

The teachers' and panel members' opinions were further compared through analysis of the correlation coefficients between the mean rating and ranking of the objectives as O-level biology course objectives, and as O-level examination objectives. The results are shown in table 7.10.

Overall the teachers and panel members views in rating and ranking of the objectives appear to coincide closely. The rank order of the teachers' and panel members' mean ratings of the course objectives was correlated with their mean order ranking of the objectives as examination objectives. A value of r = 0.5829 was obtained for teachers. The question here is what do we make of a correlation coefficient of 0.5820 and 0.5766, a striking agreement between that for teachers' (the practitioners'), and panel members' (the policy makers').

There is a slightly higher correlation between teachers' and panel members' opinions in both the rating of the course objectives according to

Table 7.06

Comparison of Teachers' and Panel members' mean rating of the course objectives according to importance in O-level biology course.

	Objective	Mean rating		ranking of the mean rating	
		Teachers	Panel members	Teachers	Panel members
	Ability to:				
1.	Show knowledge and understanding of biological facts, concepts, principles and theories.	2.878	2.750	1	2
2.	Comprehend biological texts.	2.327	2.625	12	7
3.	Apply biological principles to real life experiences.	2.878	2.750	1	2
4.	Plan experiments.	2.367	2.375	10	12
5.	Follow instructions in carrying out experimental and other biological procedures.	2.694	2.750	5	2
6.	Make and record observations accurately	2.755	3.000	3	1
7.	Interpret data to draw conclusions and make inferences.	2.714	2.625	4	7
8.	Draw biological diagrams accurately.	2.612	2.375	8	13
9.	Analyse and solve problems.	2.673	2.750	7	2
10.	Formulate hypotheses from numerical data, graphs, tables and other evidences.	2.347	2.500	11	10
11.	Explain and discuss.	2.551	2.625	9	7
12.	Write accurate reports of experiments and other biological procedures.	2.694	2.750	5	2
13.	Use manual skills to set up and operate laboratory apparatus.	2.306	2.500	13	11

Table 7.07

Comparison between teachers' and Panel members' mean ranking of the course objectives according to abilities that an O-level biology examination should test.

	Objective	Mean ranking (1-13)		ranking of the mean ranking	
		Teachers	Panel members	Teachers	Panel members
	Ability to:				
1.	Show knowledge and understanding of biological facts, concepts, principles and theories.	1.711	3.857	1	2
2.	Comprehend biological texts.	8.500	8.500	11	10
3.	Apply biological principles to real life experiences.	2.500	2.875	2	1
4.	Plan experiments.	8.304	6.000	9	6
5.	Follow instructions in carrying out experimental and other biological procedures.	5.333	4.143	4	3
6.	Make and record observations accurately.	6.000	4.625	6	4
7.	Interpret data to draw conclusions and make inferences.	5.667	5.625	5	5
8.	Draw biological diagrams accurately.	8.756	12.143	13	13
9.	Analyse and solve problems.	6.575	7.000	7	8
10.	Formulate hypotheses from numerical data, graphs, tables and other evidences.	8.400	6.833	10	7
11.	Explain and discuss.	8.605	8.333	12	9
12.	Write accurate reports of experiments and other biological procedures.	8.150	8.750	8	11
13.	Use manual skills to set up and operate laboratory apparatus.	5.271	10.375	3	12

Table 7.08

The ranking of the Teachers' and Panel members' mean rating and ranking of the course objectives according to importance in O-level biology course and according to abilities that an O-level biology examination should test.

	Objective	Ranking of teachers		Ranking of Panel members	
		mean rating	mean ranking	mean rating	mean ranking
	Ability to:				
1.	Show knowledge and understanding of biological facts, concepts, principles and theories.	1	1	2	2
2.	Comprehend biological texts.	12	11	7	10
3.	Apply biological principles to real life experiences.	1	2	2	1
4.	Plan experiments.	10	9	12	6
5.	Follow instructions in carrying out experimental and other biological procedures.	5	4	2	3
6.	Make and record observations accurately.	3	6	1	4
7.	Interpret data to draw conclusions and make inferences.	4	5	7	5
8.	Draw biological diagrams accurately.	8	13	13	13
9.	Analyse and solve problems.	7	7	2	8
10.	Formulate hypotheses from numerical data, graphs, tables and other evidences.	11	10	10	7
11.	Explain and discuss.	9	12	7	9
12.	Write accurate reports of experiments and other biological procedures.	5	8	2	11
13.	Use manual skills to set up and operate laboratory apparatus.	13	3	11	12

Table 7.09

The grand mean rating and ranking of the course objectives by Teachers and Panel members, according to importance in O-level biology; and according to abilities that an O-level biology examination should test.

	Objective	Grand mean rating (1-3)	Grand mean ranking (1-13)	Ranking of grand mean rating	Ranking of grand mean ranking
	Ability to:				
1.	Show knowledge and understanding of biological facts, concepts, principles and theories.	2.814	2.784	2	2
2.	Comprehend biological texts.	2.476	8.500	10	12
3.	Apply biological principles to real life experiences.	2.814	2.688	2	1
4.	Plan experiments.	2.371	7.152	13	7
5.	Follow instructions in carrying out experimental and other biological procedures.	2.722	4.738	4	3
6.	Make and record observations accurately.	2.878	5.313	1	4
7.	Interpret data to draw conclusions and make inferences.	2.670	5.646	7	5
8.	Draw biological diagrams accurately.	2.494	10.450	9	13
9.	Analyse and solve problems.	2.712	6.788	6	6
10.	Formulate hypotheses from numerical data, graphs, tables and other evidences.	2.424	7.617	11	8
11.	Explain and discuss.	2.588	8.469	8	10
12.	Write accurate reports of experiments and other biological procedures.	2.722	8.450	4	11
13.	Use manual skills to set up and operate laboratory apparatus.	2.403	7.823	12	9

Table 7:10

Correlations between rating of course objectives and ranking as examination objectives by teachers and panel members.

	Teachers' ranking	Panels' rating	Panels' ranking
Teachers' rating	.5829	.7430	.7206
Teachers' ranking		.5330	.6538
Panels' rating			.5766

the importance in an O-level biology course and the ranking according to the abilities that an O-level biology examination should test. The observed correlations are r = 0.7430 and r = 0.6538 respectively.

The overall rating and ranking of the objectives according to importance in O-level biology course, and according to abilities that an O-level biology examination should test correlate to the magnitude of r = 0.6253.

The above correlations have given an indication of the perceptions held by the teachers and panel members on the relative importance of the objectives as course objectives and as examination objectives. Whether these perceptions are accurate interpretations of the actual examinations weightings of the objectives is discussed in the next sections.

7.50 Content Validity of the Examinations.

There is a reasonable agreement, among different authors, such as Ebel, 1956; Lennon, 1956; Cronbach, 1970; Gronlund, 1971 , about the methods for determining content validity. Lennon's suggestion can be taken as an illustration. She says that 'content validity is evaluated by showing how the content of the test samples the class of situations or subject-matter about which conclusions are to be made' (Lennon, 1956, p.295).

According to Lennon's suggestion, and as might have been gathered from the studies on content validity reviewed in Section 7.30 of this chapter, the study of content validity involves mainly assessment by judgement rather than by the use of statistical procedures. In passing judgement on the content validity of an examination, the key question to be answered is 'to what extent does this examination require demonstration by the student of the achievements which constitute the objectives of instruction in the specified areas?' (Ebel, 1956).

The content validity of the 1983 Tanzanian biology examination has been analysed using two types of criteria:

1. in terms of coverage of the abilities intended to be developed during the four-year course.
2. in terms of coverage of the topic areas (subject-matter) intended to be covered in the four-year course.

The evidence for the two aspects of content validity was obtained through a study of the examination questions and items by eight (8) experienced examiners. The examiners' working materials and their tasks have already been described in Chapter 4, (also see Appendices IVd, IVa, and IVe), and are again summarized in the next section.

7.51 Classification of the Examination questions according to abilities being tested.

Although Bloom's Taxonomy of Educational Objectives, (Bloom, 1956) has been used extensively by researchers and examination and test designers, classification of questions into the categories of the Taxonomy has often proved to be problematic (Willmott and Hall, 1975), not least when undertaken by respondents inexperienced in the details of the Taxonomy, that is, in the six hierarchical categories of knowledge, comprehension, application, analysis, synthesis and evaluation.

Even though some researchers have reduced the number of categories by condensing the six categories into four, (Crossland and Amos, 1965; Lewis, 1967; Willmott and Nuttall, 1975; Duckworth and Hoste, 1976), their list of knowledge, comprehension, application and analysis/evaluation may still remain unclear if presented to respondents without any elaboration.

In this study, a different approach to the classification of examination questions has been employed leading to the emergence of two types of classifications:-

1. An intrinsic classification of the questions in which the O-level biology curriculum objectives and topic areas were used as criteria without reference to other systems of classification.
2. An extrinsic classification of the questions in which the intrinsic classification was related to Bloom's Taxonomy of Educational Objectives (Bloom, 1956; also see Appendix IVe).

The classification of questions according to the intrinsic criteria was done by eight experienced examiners from England and Wales. The examiners were given three tasks. After carefully studying the examination papers for 1983, (see Appendices IIIa and IIIb), the examiners were requested to:

1. Classify each examination question according to the objective which they felt the question was testing. Each item or complete question could test either a single objective or several objectives. When they felt that several objectives were being tested, then the item or question was classified on the basis of each objective tested.
2. Classify each examination question or item according to the topic areas which they felt were being tested.
3. Classify the thirteen objectives, as far as possible, according to Bloom's six categories of educational objectives. A short description of Bloom's six categories of the Cognitive Domain was supplied (see Appendix IVe).

Table 7.11 shows the total number of questions classified according to the criteria of abilities being tested by both the Theory and Practical examinations (See Appendix Va and Vb).

An item was considered to have been assigned to an ability area when at least five (5) out of eight (8) examiners agreed in their classification. There was no complete agreement between examiners in the classification, the number assigning an item to a particular ability area varied between 3 and 8. However, there was consensus agreement that all items, in addition to testing other ability areas, tested ability No.1, that is, show knowledge and understanding of biological facts, concepts, principles and theories.

Having allocated each question, item and sub-item to its respective ability area, the next step was to calculate an ability score for each of the ability areas. This was achieved by adding together the mark allocations on all the items associated with the particular ability area under consideration.

The total ability score was then considered in terms of the respective weighting accorded to each ability area in the examination. Thus, the maximum possible ability scores were used to rank the abilities according to actual examination priority. The results are shown in Table 7.12. The rankings of the abilities according to actual examination priority is discussed in section 7.54.

7.52 <u>The classification of the 13 objectives, and questions according to Bloom's Taxonomy of the Educational Objectives of the Cognitive Domain.</u>

The classification of the 13 objectives and questions into Bloom's Taxonomy of the Cognitive Domain was arrived at through two separate steps. First, the eight examiners classified, as far as possible, the 13 ability areas into Bloom's six categories of cognitive educational objectives, using the short description of Bloom's Taxonomy which was supplied (See Appendix IVe). The detailed distribution of the examiners' tallies within Bloom's six

Table 7.11

The classification of the questions according to the Objectives criteria showing the number of questions in the Theory and Practical examinations testing the different ability areas specified in the O-level biology curriculum.

	Objective	No.of items in the Theory Exam.	No.of items in the Practical Exam.	Total number of items
	Ability to:			
1.	Show knowledge and understanding of biological facts, concepts, principles and theories.	75	19	94
2.	Comprehend biological texts.	19	0	19
3.	Apply biological principles to real life experiences.	13	0	13
4.	Plan experiments.	0	1	1
5.	Follow instructions in carrying out experimental and other biological procedures.	0	7	7
6.	Make and record observations accurately.	8	5	13
7.	Interpret data to draw conclusions and make inferences.	16	6	22
8.	Draw biological diagrams accurately.	0	1	1
9.	Analyse and solve problems.	10	1	11
10.	Formulate hypotheses from numerical data, graphs, tables and other evidences.	1	0	1
11.	Explain and discuss.	45	0	45
12.	Write accurate reports of experiments and other biological procedures.	0	1	1
13.	Use manual skills to set up and operate laboratory apparatus.	0	3	3

Table 7.12

The Distribution of items and scores, and ranking of the ability areas on the basis of the weight of maximum possible scores allocated from the combined Theory and Practical examinations.

	Objective	Total No. of items from theory & practical	Maximum possible ability score	Ranking of abilities on the basis of maximum possible score
	Ability to:			
1.	Show knowledge and understanding of biological facts, concepts, principles and theories.	94	200.00	1
2.	Comprehend biological texts.	19	26.33	11
3.	Apply biological principles to real life experiences.	13	44.67	7
4.	Plan experiments.	1	35.00	9
5.	Follow instructions in carrying out experimental and other biological procedures.	7	45.00	6
6.	Make and record observations accurately.	13	75.00	4
7.	Interpret data to draw conclusions and make inferences.	22	82.50	3
8.	Draw biological diagrams accurately	1	7.00	12
9.	Analyse and solve problems.	11	40.50	8
10.	Formulate hypotheses from numerical data, graphs, tables and other evidences.	1	2.00	13
11.	Explain and discuss.	45	85.00	2
12.	Write accurate reports of experiments and other biological procedures.	1	35.00	9
13.	Use manual skills to set up and operate laboratory apparatus.	3	51.00	5

categories for the 13 objectives is shown in Table 7.13. A summary of the examiners' distribution of tallies, showing only the highest number of tallies assigned to each of the 13 ability areas in the six categories of Bloom's Taxonomy is shown in table 7.14.

Secondly, all the questions which were classified under a specific ability area in the previous section, were then ascribed, by the author to the category of Bloom's Taxonomy under which that particular ability area was assigned highest tallies by the examiners.

Table 7.15, and Table 7.16 show the total number of questions classified into each of the six categories of Bloom's Taxonomy of Educational Objectives in the Cognitive Domain for the Theory and Practical examinations respectively. Table 7.17 shows the distribution of the questions for both the Theory and Practical examinations combined. The weighting accorded to the six categories is discussed in Section 7.54.

7.53 The Classification of questions according to Topic Areas of the Syllabus.

Table 7.18 shows the total number of items assigned by the examiners to the various topic areas of the syllabus. (Also see Appendix VIa and Appendix VIb).

Once each question, item and sub-item had been categorised into its respective topic area, the next step was to calculate a topic score for each of the 16 topic areas. This was achieved by adding together the scores on all the items associated with the particular topic under consideration. The total topic score summarised the respective weighting accorded to each topic area in the examination. Thus, the total topic scores was used as a criterion for ranking the topic areas according to examinations priorities. The results are shown in Table 7.19.

Table 7.13.

The distribution of examiners tallies in classifying the 13 ability areas into Bloom's six categories of Educational Objectives of the Cognitive Domain (n = 8)

	Objective	Category of Educational Objective According to Bloom's Taxonomy of the Cognitive Domain					
		Knowledge	Comprehension	Application	Analysis	Synthesis	Evaluation
	Ability to:						
1.	Show knowledge and understanding of biological facts, concepts, principles and theories.	7	6	1	1		2
2.	Comprehend biological texts.	3	8	1	1		2
3.	Apply biological principles to real life experiences.	3	2	7			1
4.	Plan experiments.	1	2	5	1	6	1
5.	Follow instructions in carrying out experimental and other biological procedures.		2	5			
6.	Make and record observations accurately.		1	2	4	1	1
7.	Interpret data to draw conclusions and make inferences.		3	2	5	3	7
8.	Draw biological diagrams accurately	2		2	3		
9.	Analyse and solve problems.	1	3	5	8	2	4
10.	Formulate hypotheses from numerical date, graphs, tables and other evidences.	1	2	3	5	7	4
11.	Explain and discuss.	3	6	1	1	1	5
12.	Write accurate reports of experiments and other biological procedures.	1	2	1	3	6	2
13.	Use manual skills to set up and operate laboratory apparatus.	2		5			

Table 7.14

A summary of examiners' distribution of tallies showing only the highest number of tallies assigned to each of the 13 ability areas in Bloom's Categories of Cognitive Domain (n = 8)

	Objective	Category of Educational Objective According to Bloom's Taxonomy of the Cognitive Domain.					
		Knowledge	Comprehension	Application	Analysis	Synthesis	Evaluation
	Ability to:						
1.	Show knowledge and understanding of biological facts, concepts, principles and theories.	7					
2.	Comprehend biological texts.		8				
3.	Apply biological principles to real life experiences			7			
4.	Plan experiments.			6			
5.	Follow instructions in carrying out experimental and other biological procedures.			5			
6.	Make and record observations accurately.				4		
7.	Interpret data to draw conclusions and make inferences.						7
8.	Draw biological diagrams accurately.				3		
9.	Analyse and solve problems.				8		
10.	Formulate hypotheses from numerical data, graphs, tables and other evidences.					7	
11.	Explain and discuss.			6			
12.	Write accurate reports of experiments and other biological procedures.					6	
13.	Use manual skills to set up and operate laboratory apparatus.				5		

Table 7.15

Number of items in the Theory paper classified under the six categories of Bloom's Taxonomy of Cognitive Domain. (n = 75).

	Objective	Category of Educational Objective According to Bloom's Taxonomy of the Cognitive Domain					
		Knowledge	Comprehension	Application	Analysis	Synthesis	Evaluation
	Ability to:						
1.	Show knowledge and understanding of biological facts, concepts, principles and theories.	75					
2.	Comprehend biological texts.		19				
3.	Apply biological principles to real life experiences.			13			
4.	Plan experiments.			0			
5.	Follow instructions in carrying out experimental and other biological procedures.			0			
6.	Make and record observations accurately.				8		
7.	Interpret data to draw conclusions and make inferences.						16
8.	Draw biological diagrams accurately.				0		
9.	Analyse and solve problems.				10		
10.	Formulate hypotheses from numerical data, graphs, tables and other evidences.					1	
11.	Explain and discuss.		45				
12.	Write accurate reports of experiments and other biological procedures.					0	
13.	Use manual skills to set up and operate laboratory apparatus.			0			
	TOTAL	75	64	13	18	1	16

Table 7.16

Number of items in the Practical paper classified under the six categories of Bloom's Taxonomy of Cognitive Domain (n = 19)

	Objective	Category of Educational Objective According to Bloom's Taxonomy of the Cognitive Domain					
		Know-ledge	Compre-hension	Appli-cation	Analysis	Synthesis	Eval-uation
	Ability to:						
1.	Show knowledge and understanding of biological facts, concepts, principles and theories.	19					
2.	Comprehend biological texts.		0				
3.	Apply biological principles to real life experiences.			0			
4.	Plan experiments.			1			
5.	Follow instructions in carrying out experimental and other biological procedures.			7			
6.	Make and record observations accurately.				5		
7.	Interpret data to draw conclusions and make inferences.						6
8.	Draw biological diagrams accurately.				1		
9.	Analyse and solve problems.				1		
10.	Formulate hypotheses from numerical data, graphs, tables and other evidences.					0	
11.	Explain and discuss.		0				
12.	Write accurate reports of experiments and other biological procedures.					1	
13.	Use manual skills to set up and operate laboratory apparatus.				3		
	TOTAL	19	0	8	10	1	6

Table 7.17

Total number of items in both the Theory and Practical examinations classified under the Six Categories of Bloom's Taxonomy of the Cognitive Domain (n = 94)

	Objective.	Category of Educational Objective According to Bloom's Taxonomy of the Cognitive Domain					
		Knowledge	Comprehension	Application	Analysis	Synthesis	Evaluation
	Ability to:						
1.	Show knowledge and understanding of biological facts, concepts, principles and theories.	94					
2.	Comprehend biological texts.		19				
3.	Apply biological principles to real life experiments.			13			
4.	Plan experiments.			1			
5.	Follow instructions in carrying out experimental and other biological procedures.			7			
6.	Make and record observations accurately.				13		
7.	Interpret data to draw conclusions and make inferences.						22
8.	Draw biological diagrams accurately.				1		
9.	Analyse and solve problems.				11		
10.	Formulate hypothesis from numerical data, graphs, tables and other evidences.					1	
11.	Explain and discuss.		45				
12.	Write accurate reports of experiments and other biological procedures.					1	
13.	Use manual skills to set up and operate laboratory apparatus.				3		
	TOTAL	94	64	21	28	2	22

Table 7.18

The classification of the questions according to the topic areas showing the number of questions in the Theory and Practical examinations covering the different topic areas specified in the O - level biology curriculum.

	Curriculum Topic	No. of questions in the Theory Paper	No. of questions in the Practical Paper	Total No.of questions in both the Theory and Practical.
1.	Classification	11	4	15
2.	Cell structure and function	5	1	6
3.	Soil and plant growth	4	0	4
4.	Nutrition in plants and animals	15	16	31
5.	Respiration in plants and animals	9	0	9
6.	Transport system in plants and animals	15	2	17
7.	Excretion and Osmoregulation in mammals	6	2	8
8.	Control and co-ordination systems in Organisms	8	0	8
9.	Movement	2	0	2
10.	Growth and Development	7	1	8
11.	Reproduction	6	6	12
12.	Genetics	5	0	5
13	Evolution	5	0	5
14	Ecology and interdependence of organisms	6	2	8
15	Parasites and diseases of man	4	0	4
16	Man and his natural resources	2	0	2

Table 7.19

The Distribution of items and scores, and ranking of the topic areas on the basis of the weight of the total topic scores allocated from the combined Theory and Practical examinations.

	Curriculum Topic	Total No. of items from Theory and Practical	Maximum Possible Topic Score	Ranking topics on the basis of maximum possible score
1.	Classification	15	18.00	11
2.	Cell structure and function	6	8.50	12
3.	Soil and plant growth	4	21.50	8
4.	Nutrition in plants and animals	31	104.50	1
5.	Respiration in plants and animals	9	21.50	8
6.	Transport system in plants and animals	17	48.17	2
7.	Excretion and Osmoregulation in mammals	8	12.00	12
8.	Control and co-ordination systems in organisms	8	8.50	14
9.	Movement	2	1.00	16
10.	Growth and Development	8	8.00	15
11.	Reproduction	12	34.00	3
12.	Genetics	5	22.00	7
13	Evolution	5	23.00	5
14.	Ecology and interdependence of organisms	8	25.67	4
15.	Parasites and diseases of man	4	22.50	6
16.	Man and his natural resources	2	20.00	10

7.54 Discussion.

According to the examiners' classification of the questions according to abilities being tested, Table 7.11 shows that five ability areas, out of the 13 ability areas, were not tested in the Theory examination. These are ability to:-

4. Plan experiments.
5. Follow instructions in carrying out experimental and other biological procedures.
8. Draw biological diagrams accurately.

12. Write accurate reports of experiments and other biological procedures.

13. Use manual skills to set up and operate laboratory apparatus.

However, the above abilities were covered in the practical examinations. The practical examination did not test ability to:-

2. Comprehend biological texts.

3. Apply biological principles to real life experiences.

10. Formulate hypotheses from numerical data, graphs, tables and other evidences.

11. Explain and discuss.

It is clear from Table 7.11 that all the ability areas, (course objectives), were examined through the two examination papers, although to differing degrees. Again in order to compare teachers' (the practitioners'), and panel members' (the policy makers') perception of importance of the objectives as course objectives and as examination objectives with the actual examination priorities, the teachers' and panel members' mean rating and ranking of the objectives (Table 7.08), and grand mean rating and ranking of the objectives, (Table 7.09), were correlated with the ranking of the objectives by actual examinations weighting (Table 7.12). The correlation coefficients which were calculated are shown in Table 7.20.

Table 7.20.

Correlations between teachers' and panel members' rating of course objectives and ranking as examination objectives and the actual weighting of the objectives in the examination.

	Teachers' rating as course objectives	Teachers' ranking as examination objectives	Panel members' rating as course objectives	Panel members' ranking as examination objectives	Grand mean rating by Teachers + Panel as course objectives	Grand mean ranking as examination objectives
Actual Examination Weighting	.4882	.5740	.4276	.4458	.4667	.5292

Generally, the correlation values shown in table 7.20 can be described as low. When Table 7.08 and Table 7.12 are examined in more detail, obvious mismatches between teachers' and panel members' ratings as course objectives and ranking as examination objectives, and the actual ranking by examination mark weighting can be seen. These mismatches, which could be the cause of the low correlations are for example, objective No.3 (apply biological principles to real life experiences), which was ranked 1st and 2nd by teachers and panel members as course objectives, while in the actual examinations it was ranked 7th. Objective No.11 (explain and discuss), which was ranked low by both teachers (12th) and panel members (9th), as an examination objective was ranked very high (2nd) by the actual examination weighting. Other mismatches could be observed for objective No.10 (ability to formulate hypotheses from numerical data, graphs, tables and other evidences), which was ranked 7th by panel members, and yet 13th by actual examination weighting. Also objective No.13 (ability to use manual skills to set up and operate laboratory apparatus), was ranked low (12th) by panel members, but relatively high (5th) by actual examination weighting.

However, the correlations between the teachers' rating as course objectives and ranking as examination objectives in each case with actual examinations weighting are higher than those of panel members with actual examination weighting. These findings can be said to indicate that teachers stress what 'is principally designed to meet the requirements of the examinations' (Duckworth and Hoste, 1976).

Therefore, with regards to objectives coverage, on the bases of teachers' and panel members' perception of importance of the objectives as examination objectives (Table 7.09, compared with the actual examination weighting (Table 7.12), it can be said that while some of the objectives were accorded the expected weights which were thought to be desirable by teachers and panel members, (objectives No.1, No.5, No.6 and No.7), some were underweighted (objectives No.3 and No.10) and some overweighted (objectives No.11 and No.13). A more detailed consideration of practical objectives is given in Sections 8.40, 8.71 and 8.72.

With regard to the classification of the 13 objectives into Bloom's Taxonomy of Educational objectives in the Cognitive Domain, (Table 7.13), it can be seen that the examiners varied in their opinions. Most of the objectives were assigned to all possible six categories of Bloom's Taxonomy. The final classification which was used to carry out the classification of the examination questions is the one given in table 7.14.

The classification of the questions shown in tables 7.15, 7.16 and 7.17 point to the questions having tested all six categories of Bloom's Taxonomy. However, the knowledge and comprehension categories appear to have been covered more extensively than the higher cognitive categories. Views on the use of Bloom's Taxonomy in studies of this nature are given in Section 10.30.

With regard to classification of questions according to topic areas examined by the Theory and Practical examination questions, Table 7.18 shows that the Theory examination covered all of the 16 topic areas, while the Practical examination only covered eight (8) out of the 16 topic areas. However, as can be seen in column three of the same Table, it can be said that, overall all the topic areas were covered in the examination although to different degrees. The topic on 'Nutrition in plants and animals' seems to have been accorded top priority (Table 7.19).

From the author's teaching experience, a considerable longer time, is devoted to teaching the topic of 'Nutrition in plants and animals', than to most of the other topic areas, however it is questionable if the topic deserves the weighting it has been accorded in the examination. On the other hand, the topics 'Movement', 'Cell structure and function', 'Control and osmo-regulation systems in organisms' seem to have been under-examined. Otherwise, in the absence of a detailed examination item specification, the coverage of the remaining 12 topic areas can be described as satisfactory.

CHAPTER EIGHT

THE CONCEPT OF CONSTRUCT VALIDITY OF EXAMINATIONS

8.10 Introduction.

As was stated in Section 7.11, construct validity tries to explain how scores on a test are related to psychological constructs. Example of such constructs include, among others, intelligence, sex and age.

According to Cronbach and Meehl a construct is 'some postulated attribute assumed to be reflected in test performance. In test validation, the attribute about which we make statements in interpreting a test is a construct' (Cronbach and Meehl, 1955).

This chapter examines the construct validity of the 1983 Tanzanian Biology Examination with the specific intention of identifying the nature of the topic areas and abilities tested and the extent to which a 'practical mode of performance' seems to occur in these examinations. That is, investigating whether 'theoretical' and 'practical' biology examinations represent separate modes of performance. To this effect, an analysis of the Theory and Practical examination scores has been carried out, and secondly, an investigation of whether the sex of candidates, often assumed to be a 'construct' is reflected in biology examination scores to the extent of producing differences in performance in the Tanzania examinations. To this effect, comparisons of girls'and boys'scores on the Theory, Practical and overall performance have been carried out.

PART ONE

8.20 Theory and Practical Biology Examinations as Tests of Separate Constructs.

School biology is largely accepted as a practical subject. There is little disagreement with the view that practical work should be an integral part of a biology course, and that there should be some measurement of the pupils' attainment in practical work. This is reflected in the existence of separate practical assessment in examination systems, such as in the British GCE and CSE examinations; the Israel Bagrut (a series of matriculation examinations taken at the 12th grade); and the Tanzanian O-level and A-level examinations, to name but a few. Where examination papers are not given, practical abilities are assessed by coursework which is coordinated by examination boards and normally moderated by them. For example, Kingdon and Hartley (1980, 1982), have outlined and reported on internal teacher assessment of candidates' practical notebooks for the University of London Advanced level biology examinations, while Sands and Forrest (1981), have outlined the scheme of internal assessment of practical work which forms part of the Joint Matriculation Boards' (JMB) Advanced biology course assessment scheme. The scheme aims at assessing candidates' practical skills in four ability areas:

1. Possession of appropriate manipulative skills.
2. Carrying out experimental investigations.
3. Carrying out observational investigations.
4. Planning investigations.

However, there are two issues, as regards practical biology which have formed the focus of discussions. First, what should be the aims of practical biology and how should practical work in biology best be assessed, and secondly, whether practical biology is a 'unique' mode of performance. If practical skill in biology is a 'unique' mode of performance, then performances on practical examinations should be different from performances in theoretical biology examinations. Studies which have examined practical assessments as a test of a 'construct', and/or even examined the concept of construct validity of biology examinations as a whole are very few. However, there are more studies which have been reported on various aspects of the assessment of practical biology, including the nature of the abilities that a practical examination in biology should aim to assess.

8.30 Studies of Practical biology as a 'construct' and the assessment of Practical biology.

Abouseif and Lee (1965), conducted a study within physics, chemistry and biology to develop and explore tests to be used in these areas to assess the achievement of practical skills in secondary schools at fifth form level. Abouseif and Lee identified six ability areas underlying success in practical work in science, and which therefore necessitated assessment through practical tests. The six ability areas were:-

1. Ability to show knowledge of specific uses of materials, objectives and apparatus and the ability to select those which are appropriate for specific purposes (Use,U).
2. Ability to identify various kinds, sizes, types and grades of apparatus and materials (name,N).
3. Ability to perform simple operations or specific steps in complex operations (performance, P).
4. Ability to convey to others a clear statement of observed facts (report,R).
5. Ability to apply knowledge of practical methods and procedures to new experiment (solving a problem,S).
6. A capacity for methodical and tidy working, eg. in setting up and using apparatus (observe,O).
   (Abouseif and Lee, 1965, p.43).

With the help of discussion with teachers, a trial of three forms of practical tests was carried out for each subject:

Type I: The simple task-approach, comprising a suitable number of short questions.

Type II: The 'wholeness' approach, asking for the completion of one or more experiments of the customary types.

Type III: A combination of the above two approaches.

To assess the validity of the trial practical tests, they compared pupil performance on the trial practicals with performance on Attainment tests(AT); a Scientific Ability Test (SAT); and Teachers' Estimates of achievement in practical work (TE).

Abouseif and Lee reported that the overall pattern of the results suggested that each type differed in importance in the assessment of each subject because the intercorrelations between scores on each type of practical test with scores on AT, SAT, and TE were of different orders for the different subjects. They accounted for these differences in terms of differences in the ability demands of the different papers. The Type I test was found to be more suitable for assessing practical abilities in biology on the basis that it showed the highest intercorrelation (0.62) with TE than with AT and SAT (0.42 and 0.44).

Another method of assessment of practical biology has been described by Head (1966). With the help of teachers, Head developed a list of abilities which it was considered could best be tested in a practical examination, rather than in a theory examination. Eight ability areas were identified as follows:

1. Ability to observe unfamiliar objects and situations, to make records of them and deductions about them. This should not be confined to morphological exercises.
2. Ability to follow instructions accurately.
3. Ability to use reference books.
4. Ability to recognize and describe familiar biological objects, e.g. a mammalian heart.
5. Ability to use relevant manual skills, such as titrating liquids, measuring linear distances accurately, timing accurately with a stop-

watch, assembling apparatus, being sure that the apparatus used is clean.

6. Ability to draw.

7. Neatness and economy in presentation of results.

8. Ability to use a biological key (this is closely related to 1 and 3).

To assess the abilities listed above, Head designed a practical examination which could be objectively marked by using a tightly structured marking scheme. Upon correlation of practical examination scores with scores on written papers, Head reported values ranging between r=0.40 and r=0.48 for the different paper components. Commenting on the low correlation indices observed, Head said that the 'low values indicate that the theory score is not a good predictor of practical score, and that the practical examination measures something which the theory examination does not' (Head, 1966, p.95).

The need to have some measurement of pupils' performance in practical work is based on the desirability of reflecting the important role of practical work in the courses they take. Hoste (1977) reminds us of the belief that children at some stage of intellectual development learn more readily by experiencing the shapes, structure, texture, smell and seeing the interrelationships of organisms and parts of organisms, than by being told about them and reading about them. Kelly and Lister have added that 'there is also the conviction that practical work involves abilities, both manual and intellectual, which are in some measure, distinct from those used in non-practical work' (Kelly and Lister, 1969, p.123). Working with the Nuffield A-level Biology Project, they devised examinations and assessment procedures which reflected closely the aims of the course, and at the same time, constituted part of the programme for evaluating the worth of the materials and approaches used. Unlike Abouseif and Lee (1965), and Head (1966), Kelly and Lister distinguished two kinds of practical abilities. A distinction was made between those aspects which can only be undertaken in a laboratory or field situation, namely, laboratory procedures, recording and handling results. Secondly, they distinguished those which may be undertaken elsewhere, for instance as homework or in a classroom discussion, namely, hypothesis construction, experimental design and data analysis. They then point out that the second set of abilities could easily be tested in written examinations.

In their study, teachers' continuous assessment of the two aspects mentioned above was done through two methods. In one method teachers made an overall assessment of the pupils' work through a term. In the second method, a series of short, highly-structured practical tests were administered and marked by the teachers. Both methods were aimed at giving teachers more responsibility for the assessment of practical ability.

Kelly and Lister reported, among other things, the existence of uncertainties among teachers about 'the frame of reference for the grading' (Kelly and Lister, 1969). Teachers awarded higher grades in order to give their pupils the benefit of the doubt. They further reported that the distribution of grades varied between one school and another, and also between exercises within one school. However, the range of correlations between the overall continuous assessment and the results on the practical-type examination (+0.7 to -0.45) were higher than those of the tests related to the same examination questions (+0.53 to -0.64).

Tamir and Glassman (1970) have noted that, while laboratory centred activities were perhaps the most significant aspects of the BSCS curriculum, they were not assessed. Based on the argument that 'written questions cannot test the direct outcome of manual skills, observations of biological materials or the inferences which can be made from such observations' (Lister, 1969), they designed a practical laboratory examination for use with the BSCS biology curriculum (Israel version).

The examination administered to Israeli pupils, as part of their matriculation examination, was designed to test technical skills, problem solving abilities, and applications of experimental procedures.

Tamir and Glassman reported correlation coefficients between the grades in the test and the teachers' annual grades ranging from 0.43 to 0.85, and higher correlations, ranging from 0.59 to 0.86, between scores on a scholastic aptitude test and the grades in the practical examination. They further reported low inter-correlations among the different components of the Practical examination. They interpreted the low intercorrelations to mean that each portion of the practical examination measured different attributes, some of which were not included in the teachers' assessment.

In a progress report, Tamir and Glassman (1971) re-affirmed their findings from the practical examinations of 1970 that the practical examination as a whole, as well as each of its components, apparently measured some aspect of achievement which were hardly measured either by teachers' grade or by the paper-and-pencil-tests (theory examination).

In the 1971 practical examination, six skills were measured: Manipulation, Self-reliance, Observation, Investigation, Communication and Reasoning. The manipulative skills and self-reliance were assessed by the examiners during the examination period, while the other skills were assessed through the written answers.

The intercorrelations of mean scores in the various skills reported by Tamir and Glassman were low, with all values ranging between 0.01and 0.45. However, the correlation between manipulation and self-reliance was 0.68, and that between investigation and reasoning was 0.56.

Tamir (1972) subsequently distinguished three modes of performance in biology:-

1. The analytic mode: one which involves the ability to follow, understand and judge the soundness of scientific operations when they are reported in summary or research or in a scientific paper.
2. The constructive mode: one which involves the ability to contribute to scientific activities, actually discerning a problem or

constructing a hypothesis, or coming to a conclusion from data.

3. The practical mode: one based on the 'conviction that practical work involves abilities, both manual and intellectual, which are in some measure distinct from those used in non-practical work' (Tamir, 1972.)

Tamir points out that the analytic and constructive modes of performance can be tested through theory examinations. He argues however, that the practical mode is a 'distinct mode of performance' in that there are elements of practical work which cannot be tested in written examinations. These are manual skills, resource in dealing with a practical situation and direct inference from biological materials.

Finally Tamir points out that both practical and non-practical work must be seen as equally important, and therefore, the practical mode of performance should not be overlooked in any phase, and certainly not in assessment.

Differences in performance in practical and theory examinations have also been reported in other curriculum projects. Kelly (1971a) conducted an evaluation of the overall achievement of students in the Nuffield A-level Biology trials. Data used for the evaluation was obtained from a variety of sources, particularly examinations, teachers, students and external observers.

The achievement examination, the main data source, consisted of separate written and practical papers with different components and item types. There were six main components of the examination: Section A (coded answers to multiple choice and matching pair items); and Sections B, C, D, E and F (free response answers in continuous prose; short written answers to problem items; short written answers based on passages presented to examinees; and teacher assessment respectively).

Kelly reported mean correlation coefficients between A, B, C and D of 0.42 and between E and F, of 0.35;when correlating A, B, C and D against E and F (that is, the components of the theory examination against components of the practical examination), Kelly reported a low correlation coefficient of the magnitude 0.30.

Kelly's findings support Tamir's view that there is a distinct practical mode of performance in biology (Tamir 1972), because the mean correlation between modes (theory against practical) of 0.30 is lower than the mean correlation within components of the same modes (0.42 and 0.35).

Later, Hoste (1980), advanced some explanation for the different correlational values shown between the different components of the examinations in Kelly's evaluation studies. Hoste approached the problem by considering each of the examination components as being separate 'constructs'. He then factor analysed the scores of the examinations components. Although Kelly had pointed out that the practical work was 'an important, if not dominant variable', for the observed performance differences, Hoste refuted that there was evidence of the separate construct of 'practical' and 'theoretical' biology for those examinations.

Before his 1980 analyses, Hoste (1977), had determined the construct validity of the practical and theory papers of a CSE biology examination using a multitrait-multimethod matrix approach and by factor analysis. In producing the multitrait-multimethod matrix, he calculated the correlation coefficients between scores obtained by the different methods of measuring practical ability and between the different methods of measuring written theoretical biology. He hypothesised that high correlations between the components or papers would indicate convergent validity, while divergent validity would be demonstrated by low correlations between the components or paper test scores. (See Section 8.61).

Hoste reported higher correlations between the theory paper and theory paper components (ranging from 0.41 to 0.68 and 0.50 to 0.76), than between practical paper components and theory paper components (ranging between 0.27

and 0.57). From these findings he suggested that there was some degree of convergent validity between the theory paper and theory paper components and divergent validity between the theory paper and practical papers, and therefore some evidence of the existence of a 'Practical' and 'theoretical' constructs in biology.

However, upon factor analysis of the scores of theory and practical papers, Hoste reported that the theory and practical biology could not be identified as separate constructs. Therefore he concluded that he failed to find evidence that practical biology performance was a distinct mode of performance, in this particular examination.

Other aspects of practical work in biology reviewed below are mainly those which have looked into the aims of practical work. Lynch and Ndyetabura (1982a) examined the aims of teachers and the perceived influences of practical work according to students in Tasmanian High Schools and Matriculation Colleges. Lynch and Ndyetabura have listed ten (10) possible aims of practical work in science. These are:

1. To make observations more carefully.
2. To interpret observations in a logical way.
3. To make the theoretical parts of science clearer.
4. To make the theoretical parts of science more real and interesting.
5. To enable students to find out facts and principles themselves.
6. To give training in the skills and techniques of laboratory work.
7. To teach how to conduct laboratory experiments in an organized way.
8. To prepare students directly for final examinations.
9. To give personal interest in practical work and experimentation.
10. To encourage the study of science or related subjects further after leaving school.

(Lynch and Ndyetabura, 1982a, p.81).

Teachers and students in high schools and matriculation colleges were requested to choose from the list of 10 possible aims of practical work in science, four aims which they considered to be the most important aims of practical work in science. The responses were then compared.

Lynch and Ndyetabura observed that teachers at grade 7/8 (ages 11 - 13 years) had aims for practical work which differed from those at school certificate and matriculation level in that at grade 7/8 'give training skills and techniques', and 'make careful observation' were considered to be relatively more important than 'interpret observations logically', 'make theory more real and interesting', and 'enable students to find out for themselves', aims which were considered relatively more important at certificate level and matriculation colleges.

They also observed that students at high schools considered 'make careful observations' as being more important, while students in matriculation colleges considered 'train in skills and techniques', and 'make theory clearer', as being more important aims of practical work in science.

With regard to comparison between teachers' and students' views, they noted that while teachers found 'interpret observations logically' as the most important aim of practical work, for the students the most important aim of practical work in science was 'make careful observations'. However, both teachers and students regarded 'make theory clearer' as being relatively important.

In another study, Ndyetabura and Lynch (1982b) examined the preferences of science teachers and students with regard to assessment of practical work in Tasmanian High Schools and Matriculation Colleges. Five possible assessment procedures for assessing practical work in science were presented in a questionnaire and the science teachers and students were requested to indicate their preference regarding the method of assessment of practical work. The possible assessment procedures were:-

1. Continuous assessment (by teachers).
2. Practical examinations (by teachers).

3. Practical examinations (external).
4. Both practical examinations (external) and Practical examinations (internal).
5. All above (practical examinations by teachers, practical examinations external, continuous assessment).

Ndyetabura and Lynch reported a unanimous agreement between teachers and students, at all levels, that continuous assessment by teachers was the most preferred mode of assessment of practical work in science. Second in preference was practical examinations conducted by teachers. To conclude their report, they insist upon the need for assessment of practical work in science and that the measurement of practical work in science 'must substantially affect the students overall mark' (ibid. p.83).

Another look into aims of practical work in science is the study reported by Beatty and Woolnough (1982). They analysed the type and the aims of practical work which science teachers have for 11-13 year olds in England and Wales. With regard to aims, in one section of a teacher questionnaire, the teachers were asked to rate the importance of various aims for doing practical work on a five point scale. A mean rank was then calculated; and the relative order of importance was obtained between the aims, with a low mean rank indicating high importance. Twenty aims were listed.

Beatty and Woolnough reported a remarkable degree of agreement among the teachers. The following order of importance was reported for all respondents in the 11 - 13 age range for the 20 aims:-

1. To encourage accurate observation and description.
2. To arouse and maintain interest.
3. To promote a logical, reasoning method of thought.
4. To make phenomena more real through experience.
5. To be able to comprehend and carry out instructions.

6. To develop specific manipulative skills.
7. To develop certain disciplined attitudes.
8. To develop an ability to communicate.
9. To practise seeing problems and seeking ways to solve them.
10. To help remember facts and principles.
11. For finding facts and arriving at new principles.
12. To develop a critical attitude.
13. To develop an ability to co-operate.
14. To develop self reliance.
15. To give experience in standard techniques.
16. As a creative activity.
17. To elucidate theoretical work as an aid to comprehension.
18. To verify facts and principles already taught.
19. To indicate the industrial aspects of science.
20. To prepare the students for practical examinations.

(Beatty and Woolnough, 1982, p.25).

Beatty and Woolnough point out that those aims concerned with practical skills were generally rated most highly. However, they caution that the findings might not necessarily reflect the emphasis given to those aims in the laboratories.

From the studies reviewed above, there are two main views about practical assessment in biology which have been advanced:

1. There is the conviction that performance in practical assessment is a distinct mode of performance (although Hoste's study did not confirm this).
2. That not all abilities to be developed through a biology course can be assessed through practical examinations; and that there are different methods which can be employed for assessing practical work in biology.

The remaining part of the present chapter reports the part of the study concerned with the aims and the assessment of practical work in biology in Tanzanian Secondary Schools. First teachers' and panel members' opinions on the abilities which could best be assessed through practical examinations, are analysed. Secondly, the construct validity of the 1983 biology examinations is studied with the specific intention of identifying the nature of the abilities tested and the extent to which a practical mode of performance seems to occur in these examinations.

The performances of boys and girls have been considered separately as well as combined in the total sample because of the current international interest in differential attitudes to science and achievement of the sexes. This area is reviewed separately in Section 8.82.

8.40 Aims of Practical Examinations in Biology in Tanzania.

Included in Section V of the Teacher questionnaire, and Section III of the Panel members' questionnaire, was the list of the 13 objectives indicating skills which students could possess at the end of their four years biology course. (See Table 7.01, also Appendices IVb and IVc).

The teachers and panel members were asked to indicate, of the thirteen abilities those which they felt could best be assessed through a practical examination. The frequency distribution of tallies, and rank order of the tallies, for teachers' and panel members' opinions are shown in Table 8.01 and Table 8.02 respectively. Table 8.03 shows a comparison of ranks assigned to the abilities according to the Teachers' and Panel members' distribution of responses. Low mean ranking indicates high importance.

The abilities which can best be examined through practical examinations can be interpreted as the aims of practical work and practical examinations in biology. From Table 8.03, it can be seen that there is very close agreement between teachers and panel members on what the priority aims of practical examinations in biology should be. These are:-

Table 8.01

The Distribution of Total responses and rank order of the total responses according to Teachers' majority views on Abilities which can best be examined through Practical Examinations

	Objective	Distribution of Total responses (N = 49)	Ranking of the Total Responses
	Ability to:		
1.	Show knowledge and understanding of biological facts, concepts, principles and theories.	39	6
2.	Comprehend biological texts.	12	12
3.	Apply biological principles to real life experiences.	28	9
4.	Plan experiments.	32	7
5.	Follow instructions in carrying out experimental and other biological procedures.	43	3
6.	Make and record observations accurately.	48	1
7.	Interpret data to draw conclusions and make inferences.	41	5
8.	Draw biological diagrams accurately.	46	2
9.	Analyse and solve problems.	19	11
10.	Formulate hypotheses from numerical data, graphs, tables and other evidences.	21	10
11.	Explain and discuss.	1	13
12.	Write accurate reports of experiments and other biological procedures.	43	3
13.	Use manual skills to set up and operate laboratory apparatus.	31	8

Table 8.02

The Distribution of total responses and rank order of the Total responses according to Panel Members' majority views on Abilities which can best be examined through Practical Examinations.

	Objective	Distribution of total responses (N = 8)	Ranking of the Total responses
	Ability to:		
1.	Show knowledge and understanding of biological facts, concepts, principles and theories.	2	11
2.	Comprehend biological texts.	0	13
3.	Apply biological principles to real life experiences.	5	7
4.	Plan experiments.	7	4
5.	Follow instructions in carrying out experimental and other biological procedures.	8	1
6.	Make and record observations accurately.	8	1
7.	Interpret data to draw conclusions and make inferences.	5	7
8.	Draw biological diagrams accurately.	6	5
9.	Analyse and solve problems.	3	10
10.	Formulate hypotheses from numerical data, graphs, tables and other evidences.	5	7
11.	Explain and discuss.	1	12
12.	Write accurate reports of experiments and other biological procedures.	8	1
13.	Use manual skills to set up and operate laboratory apparatus.	6	5

Table 8.03

Comparison of Teachers' and Panel members' ranking of the abilities which can best be examined through Practical Examinations.

	Objective	Ranking of Teachers' opinion	Ranking of Panel Members' opinion	Mean ranking
	Ability to:			
1.	Show knowledge and understanding of biological facts, concepts, principles and theories.	6	11	9
2.	Comprehend biological texts.	12	13	12
3.	Apply biological principles to real life experiences.	9	7	8
4.	Plan experiments.	7	4	5
5.	Follow instructions in carrying out experimental and other biological procedures.	3	1	2
6.	Make and record observations accurately.	1	1	1
7.	Interpret data to draw conclusions and make inferences.	5	7	6
8.	Draw biological diagrams accurately.	2	5	4
9.	Analyse and solve problems.	11	10	11
10.	Formulate hypotheses from numerical data, graphs, tables and other evidences.	10	7	9
11.	Explain and discuss.	13	12	13
12.	Write accurate reports of experiments and other biological procedures.	3	1	2
13.	Use manual skills to set up and operate laboratory apparatus.	8	5	7

1. No.6 - Make and record observations accurately.
2. No.5 - Follow instructions in carrying out experimental and other biological procedures.
3. No.12 - Write accurate reports of experiments and other biological procedures.
4. No.8 - Draw biological diagrams accurately.
5. No.4 - Plan experiments.
6. No.7 - Interpret data to draw conclusions and make inferences.

Both teachers and panel members agree unanimously that objectives No.2 and No.11 are to be least considered as aims for practical examinations. That is, ability to comprehend biological texts, and ability to explain and discuss. Overall the rank order correlation of teachers' and panel members' priorities is very high, to the magnitude of $r = 0.8059$. The abilities which were examined by the practical examination have already been indicated in Table 7.11. There is a high level of agreement on the abilities which teachers and panel members feel should be examined by practical examinations and the abilities which were actually examined in the 1983 practical examination. These findings have given an indication of the concurrent validity of the practical examination, which in this case can be said to be high.

8.50 The Construct Validity of the 1983 Tanzanian O-level Biology Examinations.

Introduction.

There are two methods which have been used in establishing the construct validity of examination scores. One of the methods, described by Campbell and Fiske (1959) and which has been repeatedly used by Hoste (1977, 1980, 1982) in the analysis of biology examinations, is the multitrait-multimethod matrix approach. The other is by factor analysis (Cronbach and Meehl, 1955). Factor analysis has also been used by Hoste and by Fairbrother (1975). Both approaches have been infrequently employed on operational public examinations and to the author's knowledge, never on a Tanzanian examination.

In the multitrait-multimethod matrix approach, tests which measure the same traits should produce scores which correlate highly with one another, (i.e. show convergent validity), while scores on tests measuring different traits should have low intercorrelations (i.e. show divergent/discriminant validity). Therefore, if theoretical and practical biology examination performances are to be regarded as different constructs/traits, correlations between the test scores obtained on the two should be calculated and compared. The lower the between mean intercorrelations of the examinations and their components compared with the within mean intercorrelations, the greater is the evidence for divergent/discriminant validity.

In the factor analysis approach to construct validity, tests which measure the same constructs should load highly on the same factor; while tests measuring different constructs should load highly on different factors. Therefore, if performance in theoretical and practical biology examinations constitute separate constructs, that is, if they test different traits, then upon factor analysis, two factors should emerge, with the theory examination scores loading highly on one of the factors, and the practical examination scores loading highly on the other factor. Both lines of analysis have been performed on the 1983 Tanzanian biology examinations and are described and discussed below.

8.60 Construct Validity of the Theory and Practical Examination Scores by the Multitrait-multimethod Matrix Approach.

In the present study, the author was not only interested in finding out whether theory and practical examination performance are separate constructs, but also in finding out whether the different components of each of the two examinations were testing different traits. Thus, using the multitrait-multimethod matrix approach, correlations within the different components of each of the two examinations, and between the two sets of examination scores have been calculated using 'Sub-programme Pearson Correlations' of the SPSS (Nie et.al. 1975).

The input data consisted of the raw examination scores in the Theory and Practical biology examinations. That is, the same data which was used for analysing reliability in Chapter 6, and content validity in Chapter 7. The 815 candidates constitute approximately ten percent of the total 1983 O-level biology candidate population.

8.61 Results and Discussion.

As already pointed out, high intercorrelations between components are an indication that the components are measuring the same traits. On the other hand, low intercorrelations are an indication that the components are measuring different traits.

Table 8.04 carries the information about the inter-relationship of the Theory examination components. It shows the intercorrelations between the different components and the entire Theory examination. The intercorrelations are of different orders, suggesting that the different components are dissimilar. The true-false section has consistently low correlation values with every other component including the entire Theory examination with values ranging between 0.123 and 0.407. Although, generally, all other components have low correlations with each other, they correlate more highly with the entire Theory examination with values ranging between 0.592 and 0.937.

Table 8.05 shows the information relating to the Practical examination components. It shows the intercorrelations between the experimental investigation, observational investigation and the Practical examination as a whole. As in the Theory examination, the intercorrelations are of different orders, suggesting that the components are dissimilar. However, the experimental investigation section correlates more highly with the total Practical examination scores than the observational investigation section.

Table 8.04

A Multitrait-multimethod matrix of the Theory examination showing the intercorrelations between the Theory examinations Components.

Theory Paper Component	Multiple Choice	Matching Pair	True-false	Free-Response	Entire Theory Examination
Essay	.273	.319	.123	.323	.592
Multiple Choice		.490	.205	.589	.662
Matching Pair			.250	.659	.737
True-false				.337	.407
Free-response					.937

Table 8.05

A Multitrait-multimethod matrix for the Practical examination showing the intercorrelations between the Practical examination components.

Practical paper component	Observational Investigation	Entire Practical Examination
Experimental Investigation	.347	.936
Observational Investigation		.656

Table 8.06

A Multitrait-multimethod matrix for the Theory and Practical examinations showing the inter correlations between the Theory and Practical examination components.

	Paper components	Practical Exam. Components: Experimental Investigation	Observational Investigation	Entire Practical
Theory Exam. Components	Essay	.202	.401	.313
	Multiple Choice	.307	.493	.433
	Matching Pair	.378	.552	.512
	True-false	.105	.232	.172
	Free-response	.459	.656	.616
	Entire Theory	.455	.699	.630

The values reported are 0.936 and 0.656 respectively.

Table 8.06 carries the information about the two examinations' inter-relationships. It shows the intercorrelations between the different examination components of the two papers.

From Table 8.06 it can be seen that the intercorrelations are again of different orders, suggesting that the different paper components are dissimilar. The experimental investigation section has low correlations with every component of the Theory examination, with values ranging between 0.105 and 0.459. Again the true-false section of the Theory examination has the lowest intercorrelations with all sections of the Practical examination with values ranging between 0.105 and 0.232.

In order to determine the degree of divergent or convergent validity a within and between mean intercorrelation for each component of the Theory and Practical examinations was calculated. The within mean intercorrelation coefficient of a component is the mean of its intercorrelations with the other components of the same examination (For example, Essay with the other 5 components of the Theory examination); while the between mean intercorrelation coefficient of a component is the mean of its intercorrelations with the components of the other examination (for example Essay with the 3 components of the Practical examination).

The mean intercorrelation coefficient for each component was calculated by first transforming each obtained intercorrelation coefficient, (Table 8.04, Table 8.05 and table 8.06), into Zr-values through r-to-Zr transformation tables (McCall, 1980 p.374). The Zr-values were then averaged before transforming the mean Zr-value into mean r-value. The within and between mean inter-correlations of the Theory and Practical components, and the entire Theory and Practical examinations are shown in Table 8.06a.

All the between mean components' intercorrelations except for the observational investigation section of the practical paper, are lower than the corresponding within components' intercorrelations. These findings seem to demonstrate divergent/discriminant validity. The essay and true-false components of the Theory examination, and the experimental investigation section of the Practical examination have the lowest between mean intercorrelations. Therefore, they can be said to show a higher degree of divergence than the other components. On the other hand, the observational investigation section of the Practical examination seems to show the least degree of divergence.

Although the overall between mean components' intercorrelations are lower than the within mean components' intercorrelations, indicating that the papers and components under consideration seem to be measuring different facets of attainment, this does not mean for certain that the papers or components under consideration are measuring different facets of attainment because some of the differences, for example essay (0.034) and matching pair (0.036), are small. However, the precise nature of these facets of attainment is the focus of the next section where the paper component scores have been factor analysed.

Table 8.06a

The within and between mean intercorrelations of the Theory and Practical components and the Entire Theory and Practical Examinations.

Paper	Paper Component	Within Mean Intercorrelation	Between Mean Intercorrelation
Theory Paper	Essay	.339	.305
	Multiple Choice	.460	.410
	Matching Pair	.516	.480
	True-false	.270	.170
	Free-response	.650	.580
	Entire Theory Exam	.720	.600
Practical Paper	Experimental Investigation	.775	.324
	Observational Investigation	.518	.520
	Entire Practical	.846	.460

The mean intercorrelations have been calculated using a Table of Transformation of r to Zr. (McCall, 1980, p.374).

8.70 Construct Validity by the Factor Analysis Approach.

Data were analysed through 'Subprogramme FACTOR:Orthogonal Rotation: Varimax' of the SPSS (Nie et.al. 1975). The rotated orthogonal factor matrix was preferred to the rotated oblique factor matrix because the pattern matrix and the structure matrix are identical. Therefore the matrices are presented in a single table labelled factor matrix or factor loadings. That is, the coefficients, also called loading coefficients, in the matrix represent both regression weights and correlation coefficients. The choice of the principal components, which involves the extraction of principal components, which are defined as exact mathematical transformations of the original variables, was preferred because the method does not require any assumptions about the general structure of the variables. Furthermore, the method allows for principal components to be extracted from certain sets of highly correlated variables which could not be processed by other factoring methods if the correlation matrix cannot be inverted (Nie et.al.1975).

The same data used for determining construct validity by the multitrait-multimethod matrix approach in the previous section, constituted the input data. The factor programme was run first to extract initial principal factors, their Eigenvalues and corresponding proportions of variance. The results discussed, from nowon, are those of cases where the factors extracted have Eigenvalues greater than one (unity), indicating that the factor(s) are statistically significant.

Twenty four (24) initial principal factors with Eigenvalues greater than unity were extracted from the Theory examination, four (4) initial principal factors from the Practical examination, and two (2) initial principal factors from the Theory and Practical examination components treated together. These factors, their Eigenvalues and corresponding proportion of variance are shown in Table 8.07, Table 8.10 and Table 8.13 for the Theory, Practical and Theory and Practical components respectively.

Upon rotated factor analysis, five (5) factors with Eigenvalues greater than unity were extracted from the 24 initial principal factors for the Theory examination, two (2) factors extracted from the 4 initial principal factors for the Practical examination, and two (2) factors from the 2 initial principal factors for the Theory and Practical examination components treated together. The 5 factors for the Theory examination, the 2 factors for the Practical examination, and the 2 factors for the Theory and Practical examinations treated together, their Eigenvalues and corresponding proportion of variance are shown in Table 8.08, Table 8.11, and Table 8.14 respectively.

Finally, the Theory examination scores were analysed with N factor = 5, the Practical examination scores with N factor = 2, and the Theory and Practical examination components with N factor = 2. The loading of all the items of the Theory examination, and the entire Theory scores, on the 5 factors is shown in Table 8.09, the loading of all the items of the Practical examination, and the entire Practical scores, on the 2 factors is shown in Table 8.12, while the loading of all the Theory and Practical examination components on the 2 factors is shown in Table 8.15. Only loading values greater than 0.30 have been shown.

8.71 Results and Discussion.

Theory examination.

Some items load on more than one factor, while some do not load on any of the factors at all. Overall the factor loadings reported in Table 8.09 for the Theory examination questions and items range between 0.300 and 0.888. Items which load most highly on a particular factor are those which are most accurately assessing the aspect of performance represented by the factor (Hoste, 1982a). The loading patterns can be described in relation to three possible anticipated patterns:-

Table 8.07

Eigenvalues and corresponding variance for the 24 initial principal factors with Eigenvalues greater than unity extracted by the factor analysis of the Theory examination.

Factor	Eigenvalue	Percentage of variance	Cumulative percentage
1	11.258	14.8	14.8
2	2.973	3.9	18.7
3	2.146	2.8	21.5
4	1.741	2.3	23.8
5	1.713	2.3	26.1
6	1.660	2.2	28.3
7	1.487	2.0	30.2
8	1.442	1.9	32.1
9	1.383	1.8	34.0
10	1.341	1.8	35.7
11	1.290	1.7	37.4
12	1.268	1.7	39.1
13	1.244	1.6	40.7
14	1.207	1.6	42.3
15	1.198	1.6	43.9
16	1.169	1.5	45.4
17	1.157	1.5	46.9
18	1.124	1.5	48.4
19	1.100	1.5	49.9
20	1.089	1.4	51.3
21	1.052	1.4	52.7
22	1.043	1.4	54.1
23	1.034	1.4	55.4
24	1.011	1.3	56.8

Table 8.08

Eigenvalues and corresponding variance for the 5 factors with Eigenvalues greater than unity extracted from the 24 initial principal factors by the rotated factor analysis of the Theory examination.

Factor	Eigenvalue	Percentage of variance	Cumulative percentage
1	10.755	39.4	39.4
2	2.421	8.9	48.3
3	1.761	6.5	54.8
4	1.096	4.0	58.8
5	1.084	4.0	62.8

Table 8.09

Factor Loadings by items of the Theory Examination on varimax rotated factor matrix of the 5 factors extracted by the rotated factor analysis of the Theory Examination. Only loadings greater than 0.30 are shown.

Type of Items	Item	Factor loadings of items of the Theory Examination items on varimax rotated factor matrix				
		Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Essay	1					.472
	2					
	3		.484			
Multiple Choice	4					
	5					
	6					
	7					
	8					
	9					
	10					
	11					
	12	.340				
	13					
Matching Pair	14	.345				
	15					
	16	.441				
	17					
	18				.326	
	19	.327				
	20					
	21					
	22	.301				
	23	.381				
True-false	24					
	25					
	26					
	27					
	28					
	29					
	30					
	31					
	32					
	33					
Free-Response	34	.424				
	35		.395		.330	
	36			.888		
	37	.306		.770		
	38	.340		.645		
	39			.776		
	40					
	41		.420			
	42		.395		.316	
	43		.397			
	44	.329	.395		.321	
	45					
	46	.320				
	47		.366			
	48					
	49	.448				
	50	.557				
	51		.495			
	52		.619			
	53		.502			
	54					
	55	.649				
	56	.702				
	57	.560				
	58	.523				
	59	.439				
	60	.312				
	61	.303			.346	
	62	.300			.300	
	63	.435			.303	
	64				.304	
	65					
	66					
	67					
	68					
	69		.326			
	70					
	71	.377				
	72				.331	
	73					
	74				.351	
	75					

Table 8.10

Eigenvalues and corresponding variance for the 4 initial principal factors with Eigenvalues greater than unity extracted by the factor analysis of the Practical examination.

Factor	Eigenvalue	Percentage of variance	Cumulative Percentage
1	6.072	30.4	30.4
2	2.340	11.7	42.1
3	1.280	6.4	48.5
4	1.118	5.6	54.0

Table 8.11

Eigenvalues and corresponding variance for the 2 factors with Eigenvalues greater than unity extracted from the 4 initial principal factors by the rotated factor analysis of the Practical examination.

Factor	Eigenvalue	Percentage of variance	Cumulative Percentage
1	5.744	64.9	64.9
2	1.834	20.7	85.6

Table 8.12

Factor loading by items of the Practical Examination on varimax rotated factor matrix of the 2 factors extracted by the rotated factor analysis of the Practical Examination. Only Loadings greater than 0.30 are shown.

Type of Item	Items	Factor loadings of items of the Practical Examination items on varimax rotated factor matrix Factor 1	 Factor 2
Experimental Investigations	1	.794	
	2	.922	
	3	.953	
	4	.675	
	5	.900	
Observational Investigations	6		.375
	7		.607
	8		.349
	9		.403
	10		.427
	11		.433
	12		.492
	13		.501
	14		.337
	15		.386
	16		
	17		.446
	18		.390
	19		.418

Table 8.13

Eigenvalues and corresponding variance for the 2 initial principal factors with Eigenvaluesgreater than unity extracted by the factor analysis of the Theory and Practical examinations components.

Factor	Eigenvalue	Percentage of variance	Cumulative percentage
1	4.883	54.3	54.3
2	1.175	13.1	67.3

Table 8.14

Eigenvalues and corresponding variance for the 2 factors with Eigenvaluesgreater than unity extracted from the 2 initial principal factors by the rotated factor analysis of the Theory and Practical examination components.

Factor	Eigenvalue	Percentage of variance	Cumulative percentage
1	4.708	82.0	82.0
2	1.034	18.0	100.0

Table 8.15

Factor loadings by components of the Theory and Practical Examinations on Varimax rotated factor matrix of the 2 factors extracted by the rotated factor analysis of the Theory and Practical examination components. (Only loadings greater than 0.30 are shown.)

	Loadings on Varimax rotated factor matrix	
Component	Factor 1	Factor 2
Essay	.506	
Multiple-choice	.645	
Matching	.700	
True-false	.399	
Free-response	.838	.346
Entire theory	.940	.309
Exp'tal Investigation		.955
Observational investigation	.674	.350
Entire Practical	.387	.902

1. Grouping by question/item format, that is, essay; multiple-choice; matching pair; true-false or free-response.

2. Grouping by curriculum objectives, that is, according to the 13 curriculum objectives or Bloom's Categories of Educational Objectives to which the 13 curriculum objectives have been ascribed. (See Table 7.01 and Table 7.14).

3. Grouping by the curriculum content areas. (See Table 5.17).

Each of these possibilities is examined in turn in order to attempt an identification of the factor structure which has been extracted.

1. Loading by question/item format.

If the loading patterns are related to question/item format, there would be an emergence of five factor loading patterns, representing each of the five question/item formats, (essay, multiple-choice, matching-pair, true-false and free-response).

From Table 8.09, it can be seen that two of the essay questions/items 1 and 3) load significantly on two separate factors, factor 2 and factor 5 with loading values of 0.484 and 0.472 respectively.

Only 7 of the 30 objective items, one of multiple-choice (item 12), 6 of the matching-pair (items 14, 16, 18, 19, 22 and 23), and none of the true-false, have shown significant loadings (that is, greater than 0.300), on the 5 factors. Six of the seven items load on factor 1, and one on Factor 4, with loading values ranging between 0.301 and 0.441.

On the other hand, a total of 31 out of the 42 free-response items load on the five factors, with 17 of these items on factor 1 (items 34, 37, 38, 44, 46, 49, 50, 55 to 63, and 71); 10 items load on Factor 2 (items 35, 41 to 44, 47, 51 to 54, and 69); 4 items load on Factor 3 (items 36 to 39); and 10 items load on factor 4 (items 35, 42, 44, 61 to 64, 72 and 74). The loading values for the free-response items range between 0.300 and 0.888.

Therefore, from the loading patterns described above, it can be said that there is little evidence that the factor loading pattern is related to question/item format. That is, the type of question/item is not the basis of the identity of the factors.

2. Loading by Curriculum Objectives

Items 12, 14, 16, 19, 22 and 23 which according to the examiners' classification, (see Appendix Va), were viewed as testing ability No.1 (show knowledge and understanding of biological facts, concepts, principles and theories), load on Factor 1 together with items 37, 55, 56, 57, 58, 59, 60, 61, 62, and 63 which in addition to testing ability No.1, were unanimously classified as testing ability No.7, (interpret data to draw conclusions and make inferences), and ability No.6, (Make and record observations accurately), and items 44, 46, 49, and 50 which in addition to testing ability No.1, were unanimously classified as testing ability No.11, (Explain and discuss). Loading on Factor 1, also, is item 34 which was classified as testing ability No.3, ability No.9, and ability No.10, (Apply biological principles to real life experiences; Analyse and solve problems; and formulate hypotheses from numerical data, graphs, tables and other evidence).

Thus, from the fact that the items listed above are of different formats and test different content areas, covering a wide range of objectives, Factor 1 seems to be a 'general factor'.

A consideration of the items loading significantly on Factor 4, that is, items 74, 61, 72, 35, 18, 44, 42, 64, 63, and 62, listed according to loading magnitude, from highest to lowest, shows that except for item 18, the items were classified by the examiners as testing abilities beyond the knowledge and comprehension categories of Bloom's Taxonomy of Educational Objectives. Therefore, Factor 4 can be said to represent, some of the 'higher level of thinking' curriculum objectives, (No.3, No.6, No.7 and No.11) which were ascribed to the Application, Analysis, Evaluation and Comprehension categories of Bloom's Taxonomy of Educational objectives, by the examiners. (See Table 7.14).

The loadings on Factor 2 and Factor 3 are discussed in the next section.

3. Loading by Curriculum Content Areas.

In an evaluation of a CSE examination, Hoste (1977), listed the following 14 subject areas which were covered in the CSE biology course:

1. Classification.
2. Cell structure and function.
3. Movement.
4. Respiration.
5. Nutrition.
6. Excretion.
7. Circulation.
8. Water relations.
9. Soil.
10. Response.
11. Reproduction.
12. Genetics.
13. Ecology.
14. Applied biology and Hygiene.

(Hoste, 1977, p.171)

The 14 subject areas were then collapsed into 5 larger areas as follows:-

1. Anatomy and physiology.
2. Classification.
3. Evolution and Genetics.
4. Ecology.
5. Applied Biology.

(Hoste, 1977, p.174).

Following the example above, the 16 curriculum topic areas listed in Table 5.17, have been collapsed in a similar manner into 5 larger topic areas as follows:-

1. Classification.
2. Structure and Function or Anatomy and Physiology.
3. Genetics and Evolution.
4. Ecology and Interdependence of Organisms.
5. Applied Biology.

Of the items loading significantly on factor 2, item 3 is on the mechanism of breathing; items 41, 42, 43 and 44 are on the role played by the mucus in the tracheal system of man; ciliary muscles of the eye; auditory nerve of the ear; and haemoglobin of the blood respectively. Similarly, items 51 and 52 are on differences between urea and urine; vitreous and aqueous humour of the eye. All these items can be grouped under the topic area 'structure and function'.

However, item 47 is on sex cell (gamete), while item 69 is on recessive genes. The two items can be grouped together under genetics; and finally, item 53 is on the difference between food chain and food web which falls under the topic area of ecology.

Therefore, it can be said that Factor 2 is related to curriculum content areas, mainly, but not specifically, involving 'structure and function' content areas.

With regard to Factor 3, only four items load significantly on this factor, items 36, 37, 38 and 39, and all are concerned with glucose and, in particular, its structure and function. Therefore, it seems that all the items loading significantly on Factor 3, are testing the same content area. These fall into the 'structure and function' category too, hence each category in the collapsed classification seems to be too broad to accommodate the loading patterns. The factor-loading on Factor 2 and Factor 3, can therefore be said to indicate evidence of loading related to particular curriculum content areas being covered by the items under consideration. Factor 4 has already been discussed in the previous sub-section, which Factor 5 has only a single item (essay item 1) loading significantly on it.

The first factor extracted tends to be a general factor, that is, it tends to load significantly on every variable (Nie et.al. 1975, p.482). Thus,from the fact that the items loading on Factor 1 are of different formats and test different content areas, covering a wide range of objectives, Factor 1 is confirmed to be a 'general' factor.

Hoste has suggested three labels: 'general scholastic excellence' or 'general biological ability' or 'ability in written examinations' which could possibly be given to such a Factor, (Hoste, 1982b). In the present study, any of the three labels could be ascribed to this general factor- Factor 1. However, since the scores analysed are on biology, it has been tentatively labelled as representing 'general ability in biology'.

8.72 Practical Examination.

With regard to the Practical examination item scores, the factor loadings reported in Table 8.12 range between 0.337 and 0.933. Each of the Practical examination items, except item 16, has a significant loading on either Factor 1 or Factor 2. The five items of the experimental investigation section load significantly highly on Factor 1 with loading values ranging between 0.675 and 0.953. The rest of the items, belonging to the observational investigation section, load significantly on Factor 2 with loading values ranging between 0.337 and 0.607.

All the items of the Practical examination were of the free-response format. Since all the items do not load on a single factor, which would then have been related to the item format, the loading patterns can therefore be described in relation to one of the other two possible grouping patterns:

1. Grouping by curriculum objectives being tested, or
2. Grouping by curriculum content areas being tested.

The items loading on Factor 1, that is, items 1, 2, 3, 4 and 5, were unanimously classified by the examiners as testing ability No.7, (interpret data to draw conclusions and make inferences.) The possibility that items 1 to 5 could be loading on the basis of their content area is ruled out by the fact that items 14 to 19, which have very similar content areas (particularly item 18 and 19, which are most similar in that they are on food tests), load separately on Factor 2. See further elaboration in Table 8.12a and Table 8.12b. For that matter, Factor 1 can tentatively be said to be related to a single curriculum objective.

However, the loadings on Factor 2 are basically content related. Although the specimens involved are both animal and plant specimens, bearing in mind that the subject has not been taught as separate animal and plant biology, it can be said that all the items loading significantly on Factor 2 are related to the content area of 'structure and function' or 'anatomy and physiology'.

Alternatively, since items 1 to 5 loading significantly on Factor 1 belong to the experimental investigation section, while items 6 to 19 loading significantly on Factor 2 belong to the observational investigation section, Factor 1 could tentatively be said to be related to experimental investigation and thus could be labelled as 'ability to carry out experimental investigation', while Factor 2 could be said to be related to observational investigation and thus could tentatively be labelled as 'ability to carry out observational investigation'.

As can be seen from the previous paragraphs, it is not easy to relate the factors to curriculum objectives. The author agrees with Kelly (1972), who has pointed out that there are considerable problems in attempting to relate specific intellectual abilities to achievement in specific areas.

This difficulty of relating intellectual abilities to achievement was, in this study, further demonstrated when an attempt was made to factor analyse the scores of the candidates on items which had been classified as testing abilities No.3, No. 6, No.7 and No.9, in both the Theory and Practical examinations, (see Appendices 111a and 111b, and Appendices Va and Vb). These ability areas were selected because they were classified by the examiners as testing the higher cognitive categories of Bloom's Taxonomy, (See Table 7.14).

Although the input data was supposed to be related to the curriculum objectives, the output shown by the factors which emerged, Table 8.12a, were mainly content related, (Factor 2, Factor 3, and Factor 5), rather than objectives related, (Factor 4). On Factor 1 loaded the five items of the experimental investigation section which have already been discussed in detail in earlier paragraphs of this section.

Relating achievement scores to curriculum content areas seems much easier than relating achievement to intellectual abilities. For example, when the scores of the candidates on all the items in both Theory and the Practical examinations, which had been classified as testing the content areas of 'nutrition in plants and animals' (See Appendices 111a and 111b; Appendices VIa and VIb), were factor analysed, the factors which emerged, (Table 8.12b), were all specific content ralated factors, except Factor 1 on which Practical items 1 to 5 load. This factor cannot be only content related because items 36 to 39 of the Theory examination which are similar in content (glucose), load on separate factors (Factor 1 and Factor 2), and therefore Factor 1 and Factor 2 in Table 8.12b cannot both be content related.

The findings shown in Tables 8.09, 8.12, 8.12a and 8.12b are different from those reported by Hoste (1977). In Hoste's study, there was more support for the suggestion that the constructs underlying performances were more related to topic areas dealt with by the questions than for the suggestion that they tested single educational objectives in Bloom's Taxonomy. In the present study it has been shown that the constructs underlying performance in the Tanzanian 1983 O-level biology examinations were related both to topic areas dealt with by the items, (Factor 2 and Factor 3 in the Theory examination, and Factor 2 in the Practical examination), and to single curriculum objectives, also ascribed to Bloom's Taxonomy, (Factor 4 in the Theory examination and Factor 1

Table 8.12a

Factor loading by items of the Theory and Practical examinations which were classified as testing Abilities No.3; No. 6; No.7 and No.9. Only loadings greater than 0.30 are shown.

		Loading on varimax rotated factor matrix				
Paper	Item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Theory examination	1					
	4					
	5					
	6					
	34					
	35				.431	
	36		.927			
	37		.792			
	38		.674			
	39		.800			
	40				.300	
	55			.599		
	56			.619		
	57			.516		
	58			.473		
	59			.315	.351	
	60				.357	
	61				.558	
	62					
	63					
	72				.345	
	73					
	74				.414	
	75					
Practical examination	1	.702		.446		
	2	.888				
	3	.943				
	4	.688				
	5	.870				
	14					.310
	15			.333		
	16					
	17					.308
	18					.591
	19					.645

For the nature of the items in terms of abilities tested See Appendices Va and Vb.

Factor 1 - Objective No.7.
Factor 2 - Content - Glucose
Factor 3 - Content - Transport system (items 55 to 59)
Factor 4 - Objectives No.7, No.3 and No.6.
Factor 5 - Content - stem - tuber.

Table 8.12b

Factor loading by items of the Theory and Practical examinations which were classified as testing topic area No.4 - Nutrition in plants and animals. Only loadings greater than 0.300 are shown.

		Loading on varimax rotated factor matrix				
Paper	Items	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Theory examination	5					
	19			.346		
	30					
	32					
	33					
	34			.452		
	35			.492		
	36		.929			
	37		.801			
	38		.686			
	39		.794			
	40					
	45			.377		
	48					
	74					.300
Practical examination	1	.762				
	2	.900				
	3	.95				
	4	.686				
	5	.893				
	8					
	9			.339		
	12			.355	.344	
	14					
	15			.358		
	16					
	17			.378		
	18				.626	
	19				.656	

Factor 1 = Content - nutrition + Objective No.7
Factor 2 = Content - Glucose
Factor 3 = Content - nutrition
Factor 4 = Content - Stem - tuber
Factor 5 = Content - Transport system.

in the Practical examination). (See Tables 8.09 and 8.12). It was not possible to pose a single variable explanation for the factor structure which emerged.

8.73 The components of the Theory and Practical Examinations.

The loadings reported in Table 8.15 for the Theory and Practical examinations' components range between 0.309 and 0.955. Each of the Theory examination components and the Theory examination as a whole load significantly on Factor 1 with loading values ranging between 0.399 and 0.940. On the other hand, each of the Practical examination components and the Practical examination as a whole load significantly on Factor 2 with loading values ranging between 0.350 and 0.902.

It is interesting to note that the observational investigation section of the practical examination, which correlated reasonably highly with both the Theory and Practical examinations as a whole in the multitrait-multimethod matrix approach, also loads on both Factors with values of 0.674 and 0.350 respectively. This suggests that most of the items on this section could have been testing more 'theoretical' abilities than 'practical' abilities.

In order to investigate the loading pattern of groups of the items of the observational investigation section, an attempt was made to re-structure the observational investigation section according to groups of items on the basis of what the students were required to do, for example, carry out experiments (item 1); answer questions related to item 1 (items 2 to 5); identify specimens and answer questions on the specimens (items 6 to 11); draw given specimens and answer related questions (items 12 and 13); and answer questions on a given specimen (items 14 to 19).

The raw data were then factor analysed with the scores on the Theory components. The factor patterns which emerged are shown in Table 8.15a. As can be seen, overall, the factor loading patterns do not differ from those shown in Table 8.15. The separate groups of items of the observational investigation section, again, load significantly on both Factor 1 and Factor 2. Items 6 to 11 load highly only on the 'Theory' factor (Factor 1), while items 12 and 13, and items 14 to 19 load more highly on the 'Theory' factor than on Factor 2.

Therefore, on the basis of Table 8.15a, the component of the Practical examination which can be identified as a separate construct can only be the Experimental investigation section, (items 1 to 5). This has already been discussed in Section 8.72 . On the basis of Table 8.15a it can be said that Factor 1 distinctively identifies with the Theory examination components, while Factor 2 identifies with the experimental investigation section of the Practical examination.

When these results are considered side by side with the intercorrelations and mean intercorrelations between the Theory and Practical examination components reported in the multitrait-multimethod approach (Table 8.06 and Table 8.06a), it can be said that both the multitrait-multimethod approach and the factor analysis approach to determining construct validity have established the possibility of the existence of separate constructs of 'Theoretical' and 'Practical' biology in the examination performance and even within the "practical" examination.

These findings, unlike Hoste's findings (1977), support those reported by Abouseif and Lee (1965); Head (1966); Kelly and Lister (1969); and Tamir (1972), that there are two modes of performance in biology, a 'Theoretical mode of performance' and a 'Practical mode of performance'. These findings are further discussed in Chapter 10.

The organisational implications of carrying out practical assessment in biology as a whole, and specifically in Tanzania are discussed in the final Chapter alongside other issues. Part two of this Chapter, which follows immediately, compares the performances of girls and boys.

Table 8.15a

Factor Loading by components of the Theory and Practical examination on varimax rotated factor matrix, when the questions of the Practical examination are entered as separate components. Only loadings greater than 0.300 are shown.

Paper		Component	Factor Loadings	
			Factor 1	Factor 2
Theory Examination Components		Essay	.508	
		Multiple choice	.646	
		Matching-pair	.711	
		True-false	.387	
		Free-response	.840	.312
		Entire Theory	.938	
Practical Examination	Exp'tal Invest.	Item 1		.933
		Items 2 - 5		.832
	Observa- tional Investi- gation	Items 6 - 11	.629	
		Items 12 - 13	.392	.321
		Items 14 - 19	.481	.367
		Entire Practical	.386	.908

PART TWO

8.80 Sex Differences in Biology Achievement.

8.81 Introduction.

Sex is probably the single most important variable related to pupils' attitudes to science (Gardner 1975). Pupils' attitudes to science, with the element of sex differences in mind, has been studied extensively. Much more work has been reported on enrollment, attitudes and origin of the attitudes between boys and girls, than on academic performance by the two sexes. Ormerod and Duckworth (1975), and Maccoby and Jacklin (1975), have compiled a comprehensive review of researches which have looked into pupils' attitudes to science, and the psychology of sex differences, and it is not the intention of this study to repeat such reviews.

However, the present study looks at the performance of boys and girls in biology, a science subject which has been speculated to favour girls more than boys (Ormerod 1975, Murphy 1978). Apart from Comber and Keeves (1973), and Kelly (1981), who have reported on science curriculum and student achievement in nineteen countries, (work which unfortunately did not include an African country), little is available either from national random samples from different countries, or at different grade levels, on the differences between boys and girls in their performance in the different content areas of science subjects and on different ability areas. Some of the studies which have reported on performance in science, and in particular performance in biology by girls and boys are reviewed below.

8.82 Studies on Sex Differences in Achievement in Biology.

Almost all of the studies reviewed below are either studies in science achievement or studies in achievement as a whole, where sex difference in achievement was one of the factors studied, and achievement in biology one of the areas included in the analysis.

In a survey of science education in nineteen countries, Comber and Keeves (1973) have reported a strong preference by female students for the biological sciences in all countries surveyed. Achievement in different science subject areas and abilities was assessed by responses to multiple-choice items.

With regard to total science scores, Comber and Keeves have reported better performance by boys than girls, in all levels of secondary education for all the countries surveyed. As far as specific science subjects were concerned, they reported a much better performance by boys than girls in physics. In chemistry, the performance differed from country to country, for example, in French-speaking Belgium, they reported a higher performance by girls than boys, while in England, at the 14-year-old level, they reported little, if any, difference between boys and girls. The superior performance which boys showed over girls was less in biology than in physical sciences.

Backhouse (1974) compared examinations at the pass-fail borderline of GCE ordinary level for the June 1968 examinations of four GCE boards. Among the subjects studied were physics, chemistry, mathematics and biology. He reported higher achievement for boys in mathematics and biology than for girls; but for two boards he reported higher performance in physics for girls than for boys.

Later, Backhouse (1978) studied the comparability of standards in science subjects for two GCE A-level examining boards for 1973 and 1974 examinations in Nuffield Biology, Nuffield Chemistry and School Mathematics Project (SMP) Mathematics.

With regard to sex differences in achievement, Backhouse reported that the differences observed were in favour of boys for physics, and in favour of girls for biology and mathematics. The interpretation of the difference in performance between the sexes was said to be most likely due to social motivation rather than to genetic factors.

Murphy (1978) studied sex differences in examination performance in GCE O-levels for all boards for the June 1976 examinations in England and Wales. He reported differences both in the number of boys and girls entering for different subjects, and in the numbers obtaining the higher grades (ABC).

With regard to science subjects, Murphy reported a higher male entry and pass rates of 61%; 59%; 60% for males against 60%; 58%; and 56% for females for chemistry, physics and mathematics respectively. However, for biology he reported a higher female entry, 59% as opposed to 41%, but a higher male pass rate of 61% as opposed to 56% for girls.

In attempting to explain these differences, Murphy points out that it is difficult to relate specific intellectual abilities to achievement bearing in mind that 'sex differences in intellectual abilities can just as easily be the result of sex-role stereotyping', resulting from the influence of culturally defined 'sex-role stereotypes' (ibid. p.260).

He goes on to say that sex differences in examination performance could be due to the nature of the examination, in that the way examination questions are set can be biased due to the different forms of assessment used. He points out that it has been observed that objective tests produce relatively better male performance; maybe because objective tests do not test verbal ability, which has been shown to be displayed more by females.

Bloomfield et. al. (1979) have reported on ability and examinations at the 16+ age group in England and Wales. Since there are considerable differences between the syllabuses of the various CSE and GCE examining boards, Bloomfield et.al argue that a test designed to measure ability in a particular subject is not satisfactory on its own, to assess the ability of the 16+ age group. Rather they suggest the use of a test of general ability. The scores obtained on a test of general ability can then be compared with scores in a particular subject.

Upon comparison of scores obtained on Test 100, (a test of general ability), and examination scores in biology Bloomfield et.al. reported, for all 14 examining boards studied, higher mean scores for boys. For boys the mean scores ranged from 30.89 to 37.95 percent while for girls the mean scores ranged between 25.95 and 30.74 percent (Bloomfield et.al. 1979, p.156).

In another study by Murphy (1980) to determine sex differences in GCE examination entry statistics and success rates between 1951 and 1977, in England and Wales, a narrowing of the gap between the number of O-level and A-level entries received from females and males was reported. In terms of success rates, 'female candidates have, in almost every case, achieved a higher pass rate at both 'O' and 'A' level, the one exception to this trend being the 'A' level results of 1977 where, the male candidates achieved a higher pass rate than females' (Murphy, 1980, p.173). The higher pass rates for females at 'A' level have been explained by there being a smaller and more select group of girls who stay on at school to do 'A' levels.

Using marks obtained in A-level Geography examinations, Murphy reported that objective tests favoured males over females. Better performance by boys than girls on multiple choice items, in some of the science subjects, has been reported by Harding (1981).

Harding has also undertaken a comparison of performance in O-level Nuffield Biology, Chemistry, and Physics with performance in the same subjects in the conventional examinations of 1974 between boys and girls. She reported no significant overall differences between percentage passes for boys and girls in any of the six examinations studied.

However, when she analysed the data by subgroups of pupils, she reported that boys were more successful than girls in mixed schools; girls in girls' schools were more successful than girls in mixed schools; boys in comprehensive schools were more successful than girls in these schools.

Upon comparison of performance within separate components of each examination, she reported differences in mean scores obtained by boys and girls in some components of the examinations. In the multiple-choice components, mean mark comparisons showed sex differences in favour of boys in some of the science subjects and none in favour of girls. Generally, in the structured questions there were no marked differences between boys and girls. However, in the conventional biology examinations, where essay questions were included, Harding reported better performance by girls than boys.

Harding concluded by stating that the 'relative performance of boys and girls in science examinations appears to be influenced by the mode of assessment used: boys tend to achieve higher marks in multiple-choice papers and may be that girls do better in essay type questions. Structured questions appear to show the least sex bias' (Harding, 1981, p.198).

In the light of these research findings, it was felt to be vital to examine and compare performances by girls and boys on the 1983 Tanzanian O-level Biology examinations. The comparisons made include performance in the different ability areas; topic areas; overall performance; and performance on the different papers namely Theory and Practical examinations as a whole, and on the different components of the examinations.

8.83 Overall performance on the 1983 Tanzanian O-level Biology Examination.

The allocation of items to the different ability and topic areas has been discussed in Sections 7.51 and 7.53 of the previous chapter. Having allocated each item to its respective ability area, a score for each of the ability areas for all candidates was calculated. This was achieved by adding together the score on all the items associated with the particular ability area under consideration.

The ability scores were then used to compute ability means, and ability facility indices (AFI) for each ability area. The ability facility index (AFI), was calculated by the formula:-

$$AFI = \frac{\text{Mean ability score}}{\text{Maximum possible ability score}}$$

Table 8.16 and Table 8.17 show the ability statistics covering maximum possible ability score, mean ability score, and ability facility index, for the Theory and Practical examinations respectively. Table 8.18 shows the statistics for the Theory and Practical examination scores added together.

Similarly, having allocated each item to its respective topic area, a topic score was calculated for each of the 16 topic areas for all candidates. This was achieved by adding together the scores on all the items associated with the particular topic under consideration.

The topic scores were then used to compute topic means and topic facility indices (TFI) for each topic area. The topic facility index was calculated using the formula:-

$$\text{Topic Facility Index (TFI)} = \frac{\text{Mean Topic Score}}{\text{Maximum Possible Topic Score}}$$

Table 8.19 shows the topic statistics for the Theory and Practical examinations added together. It shows the maximum possible topic score, mean topic score, and topic facility index (TFI).

The performance by ability areas in the Theory and Practical examinations as indicated by the ability facility indices do not differ much (Table 8.16 and Table 8.17). Overall, as shown in Table 8.18, the ability facility values range between 0.120 and 0.291, with the majority, 10 out of the 13, values ranging between 0.214 and 0.291. Ability No.8, (ability to draw biological diagrams accurately), which was tested by one item only shows the highest ability facility index of 0.594.

The performance by topic areas, as indicated by the topic facility indices (TFI), shown in table 8.19 for a sample of 815 candidates does not differ from that already discussed in chapter 5, Table 5.19 for the sample of 519 candidates. The topics on which the students performed worst, as shown by the topic facility indices are:-

1. No.5 - Respiration in plants and animals.

Table 8.16

Overall performance in the Theory examination according to different Ability areas tested.

	Objective	Maximum Possible Ability Score	Mean Ability Score	Ability Facility Index (AFI)
	Ability to:			
1.	Show knowledge and understanding of biological facts, concepts, principles and theories.	100.00	21.067	.210
2.	Comprehend biological texts.	26.33	7.524	.286
3.	Apply biological principles to real life experiences.	44.67	9.743	.218
4.	Plan experiments.	-	-	-
5.	Follow instructions in carrying out experimental and other biological procedures.	-	-	-
6.	Make and record observations accurately.	23.00	6.851	.298
7.	Interpret data to draw conclusions and make inferences.	27.50	8.170	.297
8.	Draw biological diagrams accurately.	-	-	-
9.	Analyse and solve problems.	5.50	1.490	.271
10.	Formulate hypotheses from numerical data, graphs, tables and other evidences.	2.00	0.219	.120
11.	Explain and discuss.	85.00	13.817	.163
12.	Write accurate reports of experiments and other biological procedures.	-	-	-
13.	Use manual skills to set up and operate laboratory apparatus.	-	-	-

Table 8.17

Overall performance in the Practical examination according to different Ability areas tested.

	Objective	Maximum Possible Ability Score	Mean Ability Score	Ability Facility Index (AFI)
	Ability to:			
1.	Show knowledge and understanding of biological facts, concepts, principles and theories.	100.00	23.966	.240
2.	Comprehend biological texts.	-	-	-
3.	Apply biological principles to real life experiences.	-	-	-
4.	Plan experiments.	35.00	7.506	.214
5.	Follow instructions in carrying out experimental and other biological procedures.	45.00	10.810	.240
6.	Make and record observations accurately.	52.00	14.967	.288
7.	Interpret data to draw conclusions and make inferences.	55.00	10.539	.192
8.	Draw biological diagrams accurately.	7.00	4.157	.594
9.	Analyse and solve problems.	35.00	7.506	.214
10.	Formulate hypotheses from numerical data, graphs, tables and other evidences.	-	-	-
11.	Explain and discuss.	-	-	-
12.	Write accurate reports of experiments and other biological procedures	35.00	7.506	.214
13.	Use manual skills to set up and operate laboratory apparatus.	51.00	10.120	.198

Table 8.18

The Overall performance in the Theory and Practical examinations according to different Ability areas tested.

	Objective	Maximum Possible Ability Score	Mean Ability Score	Ability Facility Index (AFI)
	Ability to:			
1.	Show knowledge and understanding of biological facts, concepts, principles and theories.	200.00	45.033	.225
2.	Comprehend biological texts.	26.33	7.524	.286
3.	Apply biological principles to real life experiments.	44.67	9.743	.218
4.	Plan experiments.	35.00	7.506	.214
5.	Follow instructions in carrying out experimental and other biological experiments.	45.00	10.810	.240
6.	Make and record observations accurately.	75.00	21.818	.291
7.	Interpret data to draw conclusions and make inferences.	82.50	18.708	.227
8.	Draw biological diagrams accurately.	7.00	4.157	.594
9.	Analyse and solve problems.	40.00	8.995	.222
10.	Formulate hypotheses from numerical data, graphs, tables and other evidences.	2.00	0.219	.120
11.	Explain and discuss.	85.00	13.817	.163
12.	Write accurate reports of experiments and other biological procedures.	35.00	7.506	.214
13.	Use manual skills to set up and operate laboratory apparatus.	51.00	10.120	.198

Table 8.19

The Overall performance in the Theory and Practical examinations according to different Topic areas tested.

	Curriculum Topic	Maximum Possible Topic Score	Mean Topic Score	Topic Facility Index (TFI)
1.	Classification	18.00	5.681	.316
2.	Cell structure and function	8.50	1.526	.180
3.	Soil and plant growth	21.50	3.745	.174
4.	Nutrition in plants and animals	104.50	23.616	.226
5.	Respiration in plants and animals	21.50	1.520	.071
6.	Transport system in plants and animals	48.17	14.366	.298
7.	Excretion and Osmoregulation in mammals	12.00	3.441	.287
8.	Control and co-ordination systems in organisms	8.50	1.172	.138
9.	Movement	1.00	0.498	.498
10.	Growth and Development	8.00	2.038	.255
11.	Reproduction	34.00	4.514	.133
12.	Genetics	22.50	0.982	.044
13.	Evolution	23.00	0.947	.041
14.	Ecology and interdependence of organisms	25.67	4.772	.186
15.	Parasites and diseases of man	22.50	0.816	.036
16.	Man and his natural resources	20.00	3.367	.168

2. No. 12 - Genetics.

3. No. 13 - Evolution.

4. No. 15 - Parasites and diseases of man.

These topics have topic facility indices ranging between 0.036 and 0.071.

The rest of the topic areas show topic facility indices ranging from 0.133 to 0.316. The topic on 'Movement', which was tested by two items only has the highest topic facility index of 0.498.

8.84 Performance by the different sexes on the 1983 Tanzanian O-level Biology Examinations.

The data used to compare performance on the Theory and Practical examinations in Part one, and in the previous section of the present Chapter, constituted the input data for this part of the study.

The sample of 815 candidates was separated into two groups according to the sex of the candidates. The girls' sample was 258 candidates, while the boys' sample was 557 candidates. The ratio between girls and boys is approximately 1:2, the overall ratio of girls to boys entering for any compulsory examination subject in Tanzania (Examinations Council, 1983).

Two kinds of comparisons have been made. First a comparison of mean scores on the components of the Theory and Practical examinations, overall mean scores on the Theory and Practical examinations, mean and ability facility indices, and mean and topic facility indices.

Secondly, the girls' and boys' Theory and Practical examination scores, and components scores have been factor analysed separately following the same procedure outlined in Part One, Section 8.63.

8.85 Comparison of mean performance on the Theory and Practical examinations' components by the different sexes.

The performance by girls and boys on the different components of the Theory and Practical examinations, and the Theory and Practical examinations as a whole are shown in Table 8.20, and Table 8.22 respectively. The statistical significance of the differences between the means were tested through 'sub-programme 'T-Test of the SPSS (Nie et.al.1975).

Overall, the performance by girls and boys in the Theory examination as a whole differ only slightly. The mean differences range between 0.028 and 0.320. The differences and the sex favoured by the differences are shown in Table 8.21. The differences in the multiple-choice and true-false components were significant beyond the 5% level of significance. Better performance by boys than girls in objective tests has been reported by Murphy (1978). Harding has also reported a better performance by boys than girls in multiple-choice items. Murphy has stated that the better performance by boys could be due to the fact that objective tests do not test verbal ability, which has been shown to be displayed more by females than males.

On the other hand the overall performance by girls and boys in the Practical examination as a whole, and Practical examination components do differ remarkably. The mean differences range between 1.011 and 11.308. The differences and the sex favoured by the differences are shown in Table 8.23. In the Practical examination, the differences were all significant beyond the five percent (5%) level of significance.

8.86 <u>Comparison of performance by the different sexes on the different Ability areas tested by the Theory and Practical examinations</u>.

An ability facility index (AFI) for each of the ability areas tested by the Theory and Practical examinations was calculated for girls and boys using the formula already given in Section 8.83. The ability facility indices for the Theory and Practical examinations are shown in Table 8.24 and Table 8.25 respectively. The overall ability areas performance for the Theory and Practical examination scores added together is shown in Table 8.26.

Table 8.24 shows that the differences in the ability mean scores for girls

Table 8.20

Performance by Boys and Girls on the Theory Examination components and the Theory Examination as a whole.

Paper Component	Maximum Possible Score	GIRLS (N=258)		BOYS (N=557)	
		Mean Score	Standard deviation	Mean Score	Standard deviation
Essay	20	3.771	2.325	3.636	2.611
Multiple Choice	5	2.382	0.955	2.526	0.919
Matching Pair	5	1.876	1.135	1.904	1.118
True-false	5	2.766	0.746	2.925	0.808
Free-response	65	10.229	5.510	10.097	5.649
Entire Theory Paper	100	21.407	8.610	21.087	8.471

Table 8.21

The mean sex differences and the sex favoured by the difference in performance in the Theory examination and Theory Components

Paper Component	Mean difference	In favour of
Essay	.135	girls
Multiple Choice	.144*	boys
Matching pair	.028	boys
True-false	.159*	boys
Free-response	.132	girls
Entire Theory	.320	girls

* The difference is statistically significant beyond the 5% level of significance.

Table 8.22

Performance by Boys and Girls on the Practical Examination components and the Practical examination as a whole.

Paper Component	Maximum Possible Score	GIRLS		BOYS	
		Mean Score	Standard deviation	Mean Score	Standard deviation
Experimental Investigation	53	17.407	15.410	6.099	8.136
Observational Investigation	47	13.598	5.654	14.607	5.684
Entire Practical Paper	100	31.005	19.484	20.706	11.183

Table 8.23

The mean sex differences and the sex favoured by the difference in performance in the Practical examination and Practical Components

paper component	mean difference	In favour of
Experimental Investigation	11.308**	girls
Observational Investigation	1.009*	boys
Entire Practical	10.299**	girls

* The difference is statistically significant beyond the 5% level of significance.

** The difference is statistically significant beyond the 0.01 level of significance.

Table 8.24

The performance by different sexes in the Theory examination according to different Ability Areas tested.

	Objective.	Maximum Possible Score	Mean Ability Score		Ability Facility Index (AFI)	
			GIRLS	BOYS	GIRLS	BOYS
	Ability to:					
1.	Show knowledge and understanding of biological facts, concepts, principles and theories.	100.00	21.023	21.087	.210	.211
2.	Comprehend biological texts.	26.33	7.501	7.535	.285	.286
3.	Apply biological principles to real life experiences.	44.67	9.829	9.703	.220	.217
4.	Plan experiments.	-	-	-	-	-
5.	Follow instructions in carrying out experimental and other biological procedures.	-	-	-	-	-
6.	Make and record observations accurately.	23.00	7.248	6.667	.315	.290
7.	Interpret data to draw conclusions and make inferences.	27.50	8.686	7.930	.316	.288
8.	Draw biological diagrams accurately.	-	-	-	-	-
9.	Analyse and solve problems.	5.50	1.630	1.425	.296	.259
10.	Formulate hypotheses from numerical data, graphs, tables and other evidences.	2.00	0.269	0.196	.135	.098
11.	Explain and discuss.	85.00	14.000	13.733	.165	.162
12.	Write accurate reports of experiments and other biological procedures.	-	-	-	-	-
13.	Use manual skills to set up and operate laboratory apparatus.	-	-	-	-	-

Table 8.25

The performance by different sexes in the Practical examination according to different Ability Areas tested.

	Objective	Maximum Possible Score	Mean Ability Score		Ability Facility Index (AFI)	
			GIRLS	BOYS	GIRLS	BOYS
	Ability to:					
1.	Show knowledge and understanding of biological facts, concepts, principles and theories.	100.00	31.005	20.706	.310	.207
2.	Comprehend biological texts.	-	-	-	-	-
3.	Apply biological principles to real life experiences.	-	-	-	-	-
4.	Plan experiments.	35.00	13.798	4.591	.394	.131
5.	Follow instructions in carrying out experimental and other biological procedures.	45.00	17.264	7.820	.384	.174
6.	Make and record observations accurately.	52.00	21.271	12.047	.409	.232
7.	Interpret data to draw conclusions and make inferences.	55.00	18.322	6.934	.333	.126
8.	Draw biological diagrams accurately.	7.00	4.008	4.226	.573	.604
9.	Analyse and solve problems.	35.00	13.798	4.591	.394	.131
10.	Formulate hypotheses from numerical data, graphs, tables and other evidences.	-	-	-	-	-
11.	Explain and discuss.	-	-	-	-	-
12.	Write accurate reports of experiments and other biological procedures.	35.00	13.798	4.591	.394	.131
13.	Use manual skills to set up and operate laboratory apparatus.	51.00	16.135	7.333	.316	.144

Table 8.26

The performance by the different sexes in the Theory and Practical examinations according to different Ability Areas tested.

	Objective	Maximum Possible Score	Mean ability Score		Ability Facility Index (AFI)	
			GIRLS	BOYS	GIRLS	BOYS
	Ability to:					
1.	Show knowledge and understanding of biological facts, concepts, principles and theories.	200.00	52.029	41.793	.260	.209
2.	Comprehend biological texts.	26.33	7.501	7.535	.285	.286
3.	Apply biological principles to real life experiences.	44.67	9.829	9.703	.220	.217
4.	Plan experiments.	35.00	13.798	4.591	.394	.131
5.	Follow instructions in carrying out experimental and other biological procedures.	45.00	17.264	7.820	.384	.174
6.	Make and record observations accurately.	75.00	28.519	18.713	.380	.250
7.	Interpret data to draw conclusions and make inferences.	82.50	27.008	14.864	.327	.180
8.	Draw biological diagrams accurately.	7.00	4.008	4.226	.573	.604
9.	Analyse and solve problems.	40.00	15.428	6.015	.386	.150
10.	Formulate hypotheses from numerical data, graphs, tables and other evidences.	2.00	0.269	0.196	.135	.098
11.	Explain and discuss.	85.00	14.000	13.733	.165	.162
12.	Write accurate reports of experiments and other biological procedures.	35.00	13.798	4.591	.394	.131
13.	Use manual skills to set up and operate laboratory apparatus.	51.00	16.135	7.333	.316	.144

and boys for all the ability areas tested in the Theory examination are small. The ability facility indices for girls range from 0.135 to 0.316, while the ability facility indices for boys range from 0.098 to 0.290. When the differences between the means were tested for statistical significance using sub-programme T-test, (Nie and Hull, 1975), none of the differences was statistically significant at the 5% level of significance. Therefore, it can be said that there is no indication of the superiority of a particular sex in the ability areas tested by the Theory examination.

However, Table 8.25 shows that there are larger differences in the ability facility indices for girls and boys for some of the ability areas tested in the Practical examination. The ability facility indices for girls range from 0.310 to 0.573, while the ability facility indices for boys range from 0.126 to 0.604, with only one value above 0.300. The differences in the girls and boys mean ability scores ranging from 8.802 to 11.388 are all in favour of girls, while the difference between mean ability scores for ability No.8 is 0.218, and is in favour of boys. The differences between the ability means were tested for statistical significance, and all the differences in favour of girls were found to be statistically significant beyond the 0.01 level of significance. Therefore, a female sex superiority in the ability areas tested in the Practical examination are shown.

The results given in Table 8.26 for the Theory and Practical examinations performance combined has a similar pattern to that just discussed above with female superiority. The overall ability mean scores for girls range from 0.269 to 28.519, while those for boys range from 0.196 to 18.713. All the differences except one, are in favour of girls. Overall the mean ability differences shown in ability No.1, No.4, No.5, No.6, No.7, No.9, No.12 and No.13 are statistically significant beyond the 0.01 level of significance. Therefore, overall a superiority in the performance of girls in the different ability areas tested by the Tanzanian O-level biology examination has been demonstrated.

8.87 Comparison of performance by the different sexes on the different Topic areas tested by the Theory and Practical examinations.

A topic facility index (TFI) for each of the topic areas tested by the Theory and Practical examinations was calculated for girls and boys using the formula already given in Section 8.83.

The topic mean scores and topic facility indices for the different sexes are shown in Table 8.27.

With the exception of topic No.4, it can be said, generally that there is no difference in the topic mean scores and topic facility indices for girls and boys for all the topic areas tested. The differences between the girls' and boys' topic mean scores, excluding topic No.4, are very small ranging from 0.048 and 0.499 only. The topic mean score differences between the girls and boys were tested for statistical significance using sub-programme T-test (Nie and Hull (1975)).

The difference in mean topic scores for topic No.4, (which is 10.345), was found to be statistically significant beyond the 0.01 level of significance. None of the other differences showed statistical significance at 5% level of significance. Probably, the difference in topic No.4 scores must be due to the fact that a large proportion of the marks for this topic were awarded on the Practical examination in which girls performed relatively better than boys.

Overall, with regard to performance by topic areas, there is little sex superiority. Both girls' and boys' results indicate that the topics found most difficult are: Parasites and diseases of man; Evolution; Genetics; and Respiration in plants and animals.

8.88. Factor Analysis of the scores obtained by the different sexes.

The scores obtained by girls and boys on the Theory, Practical, and Theory and Practical examination components were factor analysed separately following the procedure outlined in Section 8.63.

The factor programme was run first to extract initial principal factors,

Table 8.27

The performance by the different sexes in the Theory and Practical examinations according to different Topic Areas tested.

	Curriculum Topic	Maximum Possible Score	Mean Topic Score		Topic Facility Index (TFI)	
			GIRLS	BOYS	GIRLS	BOYS
1.	Classification	18.00	5.714	5.666	.317	.314
2.	Cell structure and function	8.50	1.407	1.581	.166	.186
3.	Soil and plant growth	21.50	4.039	3.610	.188	.168
4.	Nutrition in plants and animals	104.50	30.686	20.341	.294	.195
5.	Respiration in plants and animals	21.50	1.186	1.674	.055	.078
6.	Transport system in plants and animals	48.17	14.025	14.524	.291	.302
7.	Excretion and Osmoregulation in mammals	12.00	3.545	3.392	.295	.283
8.	Control and co-ordination systems in organisms	8.50	1.002	1.250	.118	.147
9.	Movement	1.00	0.419	0.534	.419	.534
10.	Growth and Development	8.00	1.907	2.099	.238	.262
11.	Reproduction	34.00	4.560	4.492	.134	.132
12.	Genetics	22.50	0.899	1.021	.040	.045
13.	Evolution	23.00	1.025	0.910	.045	.040
14.	Ecology and interdependence of organisms	25.67	4.818	4.751	.188	.185
15.	Parasites and diseases of man	22.50	0.733	0.855	.033	.038
16.	Man and his natural resources	20.00	3.698	3.214	.185	.161

with Eigenvalues greater than unity and their corresponding proportions of the variance. Twenty three (23), and twenty six (26) initial principal factors with Eigenvalues greater than unity were extracted from the Theory examination for girls and boys scores respectively; five (5) and four (4) initial principal factors from the Practical examination for girls and boys respectively; and one (1) and two (2) initial principal factors from the Theory and Practical examinations' components for girls' and boys' scores respectively. These factors, their Eigenvalues and corresponding proportions of variance are shown in Table 8.28, Table 8.31 and Table 8.34 for the Theory, Practical and Theory and Practical examination components respectively

Upon rotated factor analysis, eight (8) factors and five (5) factors, with Eigenvalues greater than unity, were extracted from the 23 and 26 initial principal factors for the Theory examination for girls and boys respectively; two (2) and two (2) factors extracted from the 5 and 4 initial principal factors for the Practical examination for girls and boys respectively; and one (1) factor and two (2) factors from the Theory and Practical examination components for girls and boys respectively. The 8 and 5 factors for the Theory examination; 2 factors and 2 factors for the Practical examination; and 1 factor and 2 factors for the Theory and Practical examination components, their Eigenvalues and corresponding proportions of variance for girls and boys are shown in Table 8.29, Table 8.32, and Table 8.35 respectively.

Finally, the Theory scores were factor analysed with Nfactor = 8 and Nfactor = 5 for girls and boys respectively; the Practical scores with Nfactor = 2; and the Theory and Practical examination components with Nfactor = 1 and Nfactor = 2 for girls and boys respectively. The loadings on the above factors are shown in Table 8.30; Table 8.33 and Table 8.36 for the Theory examination, Practical examination and the Theory and Practical examination components respectively.

Table 8.28

Eigenvalues and corresponding variance for the 23 and 26 initial principal factors with Eigenvalues greater than unity extracted by the factor analysis of the Theory examination scores for girls and boys respectively.

GIRLS

Factor	Eigenvalue	Percentage of variance	Cumulative percentage
1	12.719	17.4	17.4
2	3.097	4.2	21.7
3	2.354	3.2	24.9
4	2.301	3.2	28.0
5	2.121	2.9	30.9
6	1.941	2.7	33.6
7	1.757	2.4	36.0
8	1.619	2.2	38.2
9	1.611	2.2	40.4
10	1.597	2.2	42.6
11	1.530	2.1	44.7
12	1.502	2.1	46.8
13	1.443	2.0	48.8
14	1.408	1.9	50.7
15	1.387	1.9	52.6
16	1.323	1.8	54.4
17	1.290	1.8	56.2
18	1.230	1.7	57.9
19	1.152	1.6	59.4
20	1.116	1.5	61.0
21	1.099	1.5	62.5
22	1.067	1.5	63.9
23	1.043	1.4	65.4

BOYS

Factor	Eigenvalue	Percentage of variance	Cumulative percentage
1	11.079	14.6	14.6
2	2.548	3.4	17.9
3	2.288	3.0	20.9
4	1.940	2.6	23.5
5	1.685	2.2	25.7
6	1.631	2.1	27.9
7	1.568	2.1	29.9
8	1.521	2.0	31.9
9	1.430	1.9	33.8
10	1.400	1.8	35.6
11	1.376	1.8	37.5
12	1.332	1.8	39.2
13	1.301	1.7	40.9
14	1.266	1.7	42.6
15	1.259	1.7	44.2
16	1.240	1.6	45.9
17	1.205	1.6	47.5
18	1.180	1.6	49.0
19	1.166	1.5	50.5
20	1.117	1.5	52.0
21	1.108	1.5	53.5
22	1.089	1.4	54.9
23	1.071	1.4	56.3
24	1.037	1.4	57.7
25	1.029	1.4	59.0
26	1.014	1.3	60.4

Table 8.29

Eigenvalues and corresponding variance for the 8 and 5 factors with Eigenvalues greater than unity extracted from the 23 and 26 initial principal factors by the rotated factor analysis of the theory examination for girls' and boys' scores respectively.

GIRLS

Factor	Eigenvalue	Percentage of variance	Cumulative percentage
1	12.278	36.2	36.2
2	2.631	7.8	44.0
3	1.865	5.5	49.5
4	1.830	5.4	54.9
5	1.604	4.7	59.6
6	1.279	3.8	63.4
7	1.095	3.2	66.6
8	1.022	3.0	69.6

BOYS

Factor	Eigenvalue	Percentage of variance	Cumulative percentage
1	10.590	36.2	36.2
2	2.212	7.6	43.7
3	1.715	5.9	49.6
4	1.338	4.6	54.2
5	1.054	3.6	57.8

Table 8.30

Factor loadings by items of the Theory Examination on Varimax rotated factor matrix of the 8 and 5 factors extracted by rotated factor analysis of the Theory examination scores for girls and boys respectively. Only loadings greater than 0.30 are shown.

		Factor loadings of items of the theory Examination items on varimax rotated factor matrix.												
	Factors	GIRLS SCORES								BOYS SCORES				
Component	Items	1	2	3	4	5	6	7	8	1	2	3	4	5
Essay	1		.370		.371								.389	
	2							.775						
	3						.775				.503			
Multiple-choice	4													
	5		.374											
	6					.300								
	7													
	8													
	9		.443											
	10													
	11									.370				
	12	.300	.317		.379					.315				
	13													
Matching-pair	14													
	15												.308	
	16			.324	.437					.424				
	17									.363				
	18		.300							.360				
	19			.412						.300				
	20				.314									
	21				.439									
	22			.422						.308				
	23				.550					.351			.306	
True-false	24						.300							.344
	25													
	26													
	27													
	28													
	29													
	30			.301										
	31													
	32													
	33		.360											
Free-response	34		.533	.326										
	35		.314					.318			.416			
	36	.850										.893		
	37	.806										.758		
	38	.719										.632		
	39	.752										.799		
	40													
	41	.402									.398			
	42					.439					.348			
	43			.345							.383			
	44			.515							.397			
	45			.560									.305	
	46			.445							.339			
	47								.376		.301			
	48										.327			
	49		.345	.320						.302	.363			
	50			.508						.356				
	51								.382		.510			
	52										.569			
	53										.448			
	54												.342	
	55	.315			.517					.657				
	56	.354	.347	.383	.485					.668				
	57		.328		.466					.549				
	58			.342	.532					.371				
	59		.472							.343				
	60						.347			.352	.362			
	61		.356							.418	.351			
	62		.451							.366				
	63		.397							.384				.325
	64					.697							.333	
	65					.700								
	66													
	67													
	68							.650						
	69							.428			.317			
	70					.808								
	71			.346										.413
	72													
	73					.603								
	74						.713			.327				
	75		.365										.339	

Table 8.31

Eigenvalues and corresponding variance for the 5 and 4 initial principal factors with Eigenvalues greater than unity extracted by the factor analysis of the Practical Examination Scores for girls and boys respectively.

GIRLS

Factor	Eigenvalue	Percentage of variance	Cumulative percentage
1	7.039	35.2	35.2
2	1.702	8.5	43.7
3	1.229	6.1	49.9
4	1.077	5.4	55.2
5	1.019	5.1	60.3

BOYS

Factor	Eigenvalue	Percentage of variance	Cumulative percentage
1	5.856	29.3	29.3
2	2.571	12.9	42.1
3	1.338	6.7	48.8
4	1.145	5.7	54.6

Table 8.32

Eigenvalues and corresponding variance for the 2 factors with Eigenvalues greater than unity extracted from the 5 and 4 initial principal factors by the rotated factor analysis as the Practical examination Scores for girls and boys respectively.

GIRLS

Factor	Eigenvalue	Percentage of variance	Cumulative percentage
1	6.725	72.0	72.0
2	1.240	13.3	85.3

BOYS

Factor	Eigenvalue	Percentage of variance	Cumulative percentage
1	5.550	61.7	61.7
2	2.080	23.1	84.8

Table 8.33

Factor loadings by items of the Practical Examination on Varimax rotated factor matrix of the 2 factors extracted by the rotated factor analysis of the Practical examination scores for both girls and boys. Only loadings greater than 0.30 are shown.

GIRLS

	Loading on varimax rotated factor matrix	
Item	Factor 1	Factor 2
1	.735	.437
2	.876	
3	.952	
4	.631	
5	.867	
6		.463
7		.703
8		
9		.305
10		.447
11		.379
12		.516
13	.353	.463
14		.323
15		.449
16		
17		.425
18		.322
19		.381

BOYS

	Loading on varimax rotated factor matrix	
Item	Factor 1	Factor 2
1	.751	
2	.935	
3	.950	
4	.745	
5	.916	
6		.350
7		.555
8		.371
9		.416
10		.398
11		.421
12		.473
13		.492
14		.348
15		.389
16		
17		.433
18		.439
19		.462

Table 8.34

Eigenvalues and corresponding variance for the 1 and 2 initial principal factors with Eigenvalue greater than unity extracted by the factor analysis of the Theory and Practical examination components for girls and boys respectively.

GIRLS

Factor	Eigenvalue	Percentage of variance	Cumulative Percentage
1	5.615	62.4	62.4

BOYS

Factor	Eigenvalue	Percentage of variance	Cumulative Percentage
1	4.726	52.5	52.5
2	1.227	13.6	66.1

Table 8.35

Eigenvalues and corresponding variance for the 1 and 2 factors with Eigenvalue greater than unity extracted from the 1 and 2 initial principal factors by the rotated factor analysis of the Theory and Practical examination components for girls and boys respectively.

GIRLS

Factor	Eigenvalue	Percentage of variance	Cumulative Percentage
1	5.305	100.00	100.00

BOYS

Factor	Eigenvalue	Percentage of variance	Cumulative Percentage
1	4.533	81.4	81.4
2	1.038	18.6	100.0

Table 8.36

Factor loadings by components of the Theory and Practical examinations on varimax rotated factor matrix for the 1 and 2 factors extracted by the rotated factor analysis of the components. Only loadings greater than 0.30 are shown.

GIRLS	Loadings on varimax rotated factor matrix
Component	
Essay	.531
Multiple-choice	.689
Matching-pair	.734
True false	.347
Free-response	.903
Entire theory	.976
Exp'tal Investigation	.817
Obse'nal Investigation	.799
Entire Practical	.903

BOYS	Loadings on varimax rotated factor matrix	
Component	Factor 1	Factor 2
Essay	.510	
Multiple-choice	.592	
Matching-pair	.711	
True false	.398	
Free-response	.822	.354
Entire theory	.944	
Exp'tal Investigation		.929
Obse'nal Investigation	.660	.381
Entire Practical	.404	.874

8.89 Results and Discussion.

The factor loadings reported in Table 8.30 for the Theory examination items range from 0.300 to 0.850 and from 0.301 to 0.893 for girls and boys respectively. For both girls and boys, the essay questions load significantly on separate factors. The loading values for the essay questions for the girls are higher than those for boys. The loading values for the essay questions are 0.370 and 0.371, 0.553 and 0.775 for girls, while those for boys are 0.389 and 0.503.

Most items, 61 out of the 75 items, for the girls' scores load on four Factors while 86 out of 75 items for the boys' scores load on three Factors. Only 14 and 10 objective items for girls' and boys' respectively show significant loadings on the extracted factors, while the loadings for the girls' scores are spread out on 6 factors, those for the boys load in two factors. For the free-response items, 35 and 33 out of the 42 items load significantly for girls' and boys' scores respectively.

The process and problems of labelling the factors extracted from the girls' and boys' scores in the Theory and Practical examinations, and the Theory and Practical examination components do not differ from those already discussed and expressed in Section 8.70 to Section 8.74.

However, a comparison of the items' loading patterns for the girls' and boys' scores, (Table 8.30), with the overall loading patterns, (Table 8.09) show that:-

Girls		Boys		Whole sample
Factor 1	=	Factor 3	=	Factor 3
Factor 4	=	Factor 1	=	Factor 1
		Factor 2	=	Factor 2

Thus, it can be said that, the boys' loading patterns are more similar to the overall loading patterns than are the girls loading patterns. The labelling or equating of the rest of the girls' factors to the boys' or the whole sample factors was further complicated by the emergence of the 8 factors as opposed to the 5 factors for the boys and overall samples.

With regard to the Practical examination items, the factor loadings reported in Table 8.33 range between 0.322 and 0.952, and 0.350 to 0.950 for girls and boys respectively. For both girls and boys, item 16 does not load significantly on either of the factors. For both girls and boys Factor 1 and Factor 2 loading patterns resemble the overall loading patterns for the Practical examination shown in Table 8.12 with experimental investigation items loading significantly on Factor 1, and observational investigation items loading significantly on Factor 2.

Perhaps the most interesting finding was that when the Theory and Practical components' scores were factor analysed, a single factor emerged from the girls' scores and two factors from the boys' scores (Table 8.36). Whether the boys' loading pattern had been influenced by the fact that there were more boys in the sample than girls was checked by factor analysing scores of 258 boys (equals girls' sample) picked alternatively using even numbers. The factors which emerged are shown in table 8.37. As can be seen the loading pattern is similar to that shown in Table 8.36 for boys. Therefore, the sample size did not seem to influence the loading patterns.

The single factor extracted from the girls scores could tentatively be labelled as 'general biology ability' (Hoste, 1982b). On the other hand, the two factors from the boys scores, support the existence of the constructs of 'Theoretical' and 'Practical' biology in examination performance because, like for the overall sample, Table 8.15, the components of the Theory examination and the Theory examination as a whole load significantly on Factor 1, while those for the Practical examination components and the Practical examination as a whole load significantly on Factor 2. Thus Factor 1 is related to 'Theoretical biology performance' while Factor 2 is related to 'Practical biology performance' - a 'distinct mode of performance' (Tamir, 1972).

The differences in the loading patterns for the Theory and Practical examinations' components for the girls and boys samples are further discussed in section 9.74.

Table 8.37

Factor loadings by components of the Theory and Practical examinations on varimax rotated factor matrix for 2 factors extracted by the rotated factor analysis of the components for 258 boys. Only loadings greater than 0.30 are shown.

	Loadings on varimax rotated factor matrix	
Component	Factor 1	Factor 2
Essay	.510	
Multiple-choice	.575	
Matching-pair	.713	
True-false	.363	
Free-response	.795	
Entire theory	.936	
Exp'tal Investigation		.926
Obse'nal Investigation	.653	.405
Entire Practical	.419	.868

CHAPTER NINE

CONTINUOUS ASSESSMENT

9.10 Definition.

Assessment of students' attainment is an integral part of any educational process, and every teacher, at whatever level he teaches is called upon to test his students. In order to function effectively, teachers require information concerning the levels of attainment achieved by their students. Methods of measuring attainment vary from personal judgements made by teachers to tests which have national or even international recognition. Thus, general course-work, practical work, oral tests, project work, and examinations can all have their places in assessment of students' attainment.

In addition to serving the normal routine of assessment of students' attainment, teacher assessment can operate on other levels. Cohen and Deale (1977) have distinguished three different views of how teacher assessments could operate. First, where the teacher is seen as acting as a counterpart to the external examiner; assessing the same skills and abilities over a period of time as are tested on the day of the external examination. The teachers' assessments are scaled and weighted appropriately and are taken into account together with the external papers in deciding the candidates' final grades.

Secondly, where the teacher is asked to assess aspects of the course which either cannot be examined externally or can be only with great difficulty. For example, conducting of oral language examinations, and assessment of field work in subjects like geography and biology.

Thirdly, where the teacher not only assesses specific components, such as oral, field work, but also marks all the externally set papers in the final examinations.

The system of continuous assessment which operates in Tanzania and is reported in this study, encompasses the ideas reflected in the first view described by Cohen and Deale, above. That is, a continuous assessment scheme (CA), is one which involves a systematic collection of marks or grades by the teacher over a period of time and the consolidation of the marks or grades into a final score taken into account in deciding the candidates' final grades. The most distinctive feature of CA in general, is that it is a cumulative process, developing as the pupil develops and reflecting his changes in response to the course (Hoste and Bloomfield, 1975).

9.20 Studies on Continuous Assessment.

Little has been reported on continuous assessment in science. As will be seen from the studies reviewed below, most continuous assessment which has been reported is related to assessment of practical work in the science subjects, replacing the traditional practical examinations.

Chalmers and Stark (1968) have described an assessment scheme prepared for the Scottish HNC course in practical chemistry. In the scheme, teachers were required to plan practical courses, based on the subject matter according to the demands of the Joint Committee for National Certificates and Diplomas, the Chemistry Joint Committee.

The assessment by teachers was based on five ability areas.

Ability to:-

1. Work accurately.
2. Use equipment effectively.
3. Calculate results correctly.
4. Write a clear, concise and accurate report.
5. Demonstrate an understanding of the principles underlying each experiment.

(Chalmers and Stark, 1968, p.155).

Marks were allocated for accuracy, calculation, methodology and comprehension in the ratio 40:20:20:20 respectively. The minimum assessments for each student had to be at least ten experiments. Each major section of the syllabus had to be covered.

Chalmers and Stark reported that teachers observed the beneficial effects of the scheme in matters such as discussion of the work, use made of the library, less careless work and on the students' skills at organizing their work. However, they also reported that both teachers and students commented on the increased work load. For example, they reported that the 'teacher has to devote much more time to organizing, supervising and marking the work, the student to the preparation of reports' (ibid. p.155).

Eggleston and Newbould (1969) have reported on measurement of attainment in CSE physics. Three measures of attainment were considered: scores on a moderating test; teachers' assessments; the CSE grade and mark in physics. Four categories of performance formed the basis for teachers' assessments and the moderating test. These were the ability to:-

1. Make observations.
2. Organize data.
3. Make inferences from evidence.
4. Apply facts and principles to problem solving.

(Eggleston and Newbould, 1969, p.75).

They observed that the intercorrelations between the four categories in the moderating test were much lower than with same categories when assessed by teachers. The intercorrelations ranged from 0.14 to 0.34, and 0.44 to 0.54 for the moderating test and teachers' assessment respectively. Commenting on the high intercorrelations between the four categories in the teachers' assessments, Eggleston and Newbould said that the high intercorrelations 'was an indication of a greater overlap between the attainments measured in each of the four categories when the teachers assessed their pupils' attainment than was the case with the four corresponding categories in the moderating test' (ibid. pp.60.61).

Upon factor analysis of the intercorrelation matrices between the assessments made by each teacher in the four categories, the factor structure which emerged indicated the existence of two groups of teachers. The first group gave a factor pattern in which three factors accounted for all or almost all the variance with factor I accounting for over 70% of the variance. The second group gave a factor pattern in which four factors emerged with Factor 1 accounting for 55% or less of the total variance.

The weight given to, and the implementation of, CA vary greatly from country to country and within individual countries. For example, Young (1970), points out that individual practices of continuous assessment in the United States vary so widely, that it is not even possible to describe its administration in detail. Young also points out that, in general, continuous assessment is based on work assigned to pupils by the teacher. The assessment then usually consists of a series of questions about the assignments which are given regularly or irregularly (weekly, daily or fortnightly).

Young states that with continuous assessment, the student is under continuous tension, never quite free from worry until the course is completed though, ideally, he keeps up with his work schedule and progresses steadily. He also indicates that continuous assessment leads to disjointed comprehension, if topics already assessed are not to appear in later assessments.

With regard to teachers, he says that 'when a teacher has to write several examinations during a year, he may tend to get a bit hurried and write poorly phrased questions which, not even he could answer satisfactorily' (Young, 1970, p.13).

In another study, Hoste and Bloomfield (1975) have compared grades awarded for one of the 1972 CSE examining boards in Yorkshire, using examination method (EM), and continuous assessment method (CAM). Using Test-100 to make comparisons between EM scores and CAM scores for 8 subjects, they reported higher correlations between Test-100 with EM scores than test-100 with

CAM scores for all the subjects, with values ranging from 0.41 to 0.64, and 0.02 to 0.57 for boys; and from 0.34 to 0.70, and from 0.16 to 0.46 for girls, for EM and CAM respectively. For biology, the correlations were 0.57 and 0.02 for boys; and 0.48 and 0.23 for girls, for EM and CAM respectively. For the boys, they reported higher correlation coefficients for EM than CAM in all subjects with the exception of physics. For the girls, however, while this was still true of mathematics and the sciences, the correlations of test score and grades in English, geography and history were similar/almost similar for both EM and CAM.

Hoste and Bloomfield, after considering estimates of severity and leniency of grades awarded for girls and boys, concluded that there was no evidence that standards were different for candidates when grades were awarded by EM and CAM.

Another scheme of teacher assessment is one which has been described by Sands and Forrest (1981). The scheme is for teacher assessment of practical work in A-level biology and its moderation for the Joint Matriculation Board (JMB). In the scheme, the practical work, is assessed throughout the two years of the A-level course. The accumulated mark awarded constitutes 20 per cent of the A-level grade.

The practical work used by teachers to assess a candidate's practical skill may be chosen from the whole syllabus. Teachers are required to supervise all assessed practical work and students are assessed on the bases of four abilities. These are:-

1. Possession of appropriate manipulative skills.
2. Carrying out observational investigations.
3. Carrying out experimental investigations.
4. Planning investigations.

(Sands and Forrest, 1981, p.147).

The teachers' assessments are moderated through the scores of a compulsory section in the written examination (Paper 1 Section A). A satisfactory correlation coefficient between teachers' assessment and scores on Paper I Section A was set to be 0.60. For the first trials, the correlation coefficients calculated ranged between 0.57 and 0.60. Thus the relationship between teacher assessments and the moderating instrument, (Paper I Section A), was considered to be of the correct order.

Sands and Forrest in conclusion commented that 'it is to be expected that the validity of an examination will be enhanced if there is an element of teacher assessment in the aggregated result'. They went on to say, 'although individual teachers can put their candidates into a valid and reliable order of merit, they cannot be expected to relate accurately the performances of their own candidates to those of other candidates in other centres. The function of a moderating process is to do this' (Sands and Forrest, 1981, p.150).

As mentioned in Chapter I, continuous assessment, in Tanzania is part of the national assessment scheme and is an essential component of O-level and A-level assessment of students' achievement. The CA score awarded to each student in the different subjects contributes 50% to the final School Certificate Examination score for each particular subject studied and examined.

The rest of the present chapter describes CA in Tanzania and then looks into some aspects of continuous assessment in biology, including a comparison of continuous assessment scores (CA scores) with the final examination scores (FE scores) awarded to candidates in O-level biology in 1983.

9.30 Continuous Assessment in Tanzania.

9.31 General Description of the Continuous Assessment Scheme.

As was mentioned in Chapter I, continuous assessment has been an integral component of the student assessment scheme in all Tanzanian O-level and A-level subjects since 1976.

The main purpose of having a continuous assessment scheme as an integral component of assessment procedures in the Tanzanian education system is to eliminate/minimize the element of risk associated with a single examination, and to give a valid indication of student achievement, because it is felt that no student who works conscientiously should fail.

By the Musoma Resolution on the implementation of Education for Self-Reliance (ESR), it was therefore resolved that continuous assessment be an integral component of assessment of student achievement:-

> 'We have to get rid of the ambush type of examination. At the moment we are placing too much emphasis on written examinations..........the excessive emphasis placed on written examinations must be reduced and that students' progress in the classroom, plus his performance of other functions and the work which he will do as part of his education, must all be continuously assessed and the combined result is what should constitute his success or failure'. (TANU, 1974, p.21).

Continuous assessment score (CA score) contributes 50% to the total weight in every subject examined. Continuous assessment is implemented with the assumption that the student obtains a mark which reflects fairly his ability and also shows him where he needs to improve his work before the final assessment by the end-of-course examinations.

The implementation of continuous assessment is the responsibility of teachers in the schools. Subject teachers are required, by the scheme, to keep an accurate record of students' academic performance in the last two years of their four year O-level education, and throughout the two years of A-level education.

The performance is recorded on special forms supplied to all schools by the Examinations Council of Tanzania. There are three component scores to be recorded for each student for each subject studied and examined. These are: Exercises, tests, and projects.

Exercises.

The Examinations Council requires an exercise to be given after every six taught periods for O-level subjects, and every eight periods for A-level subjects. Most O-level subjects have 3-4, forty-minute, taught periods per week. Thus, teachers are required to give an exercise approximately every two weeks.

Tests.

Tests are required to be administered under examination conditions. Teachers can give tests as often as they desire, covering any part of the syllabus already taught. However, the Examinations Council requires only the records of the tests administered at the end of every school term. There are two school terms each year. Thus the teachers have to record two tests for each student, for each subject every school year.

Projects.

Projects are started in Form 3 (third year of secondary school), and are completed while in Form 4. There are five subject groups from which students can choose projects:-

1. Arts and Social Sciences.
2. Languages.
3. Science or Technical subjects.
4. Agriculture or Business.
5. Home Economics or Military Science.

(Mtani, 1983, p.12).

The teachers within a subject group are responsible for setting a number of projects and for formulating the guidelines which students should use to do their selected projects. Students are required to choose their specific projects from the list of projects formulated by the teachers. All projects are done in groups not exceeding five members per group. The projects are marked by project marking panels which are wholly organized by the teachers.

The final distribution of marks within the three CA components is in the ration 25:20:5 for exercises, tests and projects respectively.

9.32 Moderation of Continuous Assessment Scores.

Mtani (1983) has pointed out three things which make standardisation of CA scores a necessity. First, the fact that teachers are given the opportunity to set exercises and tests using purely their own judgement to determine the standards of the items used for assessment, leads to varying standards of items used from one school to another school. Secondly, following directly from the first point, the marking schemes are bound to vary, thus giving rise to large disparities in the students' marks in different schools for the different subjects. Thirdly, leniency is shown by teachers to their own students, thus making CA scores, on the whole, much higher than the final examinations scores.

Therefore, the final examination, for each subject is used as the moderating instrument for the CA scores for that particular subject. The general procedure for standardisation has been described by Mtani (1983) as involving six main steps. These steps involve, for each subject, the calculation of:-

1. National Mean for Continuous Assessment ($\overline{NCA}$).
2. National Mean for Final Examination ($\overline{NFE}$).
3. National Inflation of Continuous Assessment mark, also called National Mean Difference (N.M.D).
   N.M.D. = $\overline{NCA} - \overline{NFE}$.
4. Centre mean for Continuous Assessment ($\overline{CA}$), and Centre mean for Final Examinations ($\overline{FE}$).
5. Centre Inflation of Continuous Assessment mark also called Centre Mean Difference (M.D).
   M.D = $\overline{CA} - \overline{FE}$.

6. Adjustment Factor (A.F,) for each centre. That is, the difference between the National Inflation (N.M.D) and the Centre Inflation (M.D) of continuous assessment mark.
   A.F. = N.M.D - M.D.

The adjustment factor is the number of marks which have to be added or subtracted from each student's marks in the centre for a particular subject depending on whether the A.F. value is positive or negative. A negative A.F value indicates a measure of leniency exercised by the teachers, and a positive A.F value indicates a measure of severity exercised by the teachers.

9.40 Methods Used by Teachers to Obtain a Mark for CA.

It was within the interest of this study to find out how teachers awarded marks for the exercises, tests and projects. Two things were of interest: first the method used to decide on the exercises and tests to be recorded; and secondly, whether the actual marks were awarded by impression mark or mark scheme.

The teachers were presented, in Section Six of the Teacher questionnaire, with a list of six possible methods which could be used for obtaining a continuous assessment mark for biology. The six possible methods were identified by the author, through experience, from a range of possible methods used in schools in Tanzania. The teachers were required to indicate their order of use of the methods by ranking them from 1 (most used) to 6 (least used).

The teachers' frequency of ranking of the six methods, including the mean ranks, and ranking of the mean ranks is shown in Table 9.01.

The stability of the teachers' responses was checked from data derived through interviews given to 19 teachers who had previously completely the questionnaire and who were attending a biology workshop during the time of the study (See Section 4.60). A comparison of the interview responses with the questionnaire responses of the 19 teachers interviewed showed an exactly similar pattern of responses.

From Table 9.01, it can be said that the two most frequently used methods for obtaining a continuous assessment mark are: class tests specially prepared for this purpose and classwork exercises of ongoing work. When the teachers who were interviewed were asked whether they should use the classwork exercises of ongoing work, or class tests specially prepared only, they were of the opinion that either method was appropriate provided the tests and exercises were based on the work which formed part of the normal teaching. However, they complained that the weight given to project work (10% of the total CA weight), is too little compared to the time and amount of work involved in organizing, carrying out and marking the project work.

Table 9.01

Frequency of ranking of methods used for obtaining a continuous assessment mark for biology.

Method of assessment	frequency of ranking of method 1 to 6						Mean rank	ranking of the mean rank
	1	2	3	4	5	6		
classwork exercise of ongoing work	22	8	7	9	1	1	2.163	2
class tests specially prepared	20	8	16	4	0	0	2.041	1
practical exercise of ongoing work	2	20	6	13	6	1	3.020	3
practical test specially prepared	1	8	9	17	11	2	3.653	4
term examination	3	4	8	4	23	6	4.122	5
project work or special studies	0	0	2	1	7	38	5.571	6

9.50 The Award of Marks.

The information on the kind of abilities and skills to be assessed, and whether the actual marks were awarded by impression mark or mark scheme was mainly gathered from the 19 teachers who were interviewed during the workshop.

From the interviews, and from the general description of the continuous assessment scheme, (Section 9.31 to Section 9.34), it was clear that, up to the time of the study, there were no formal guidelines from the Examinations Council which provided the teachers with the criteria for assessment, that is, which told the teachers the abilities and skills to assess.

From the interviews, it was clearly indicated by the teachers that marks are awarded both by impression mark and mark scheme. Impression marking was more frequently used to award marks to classwork exercise of ongoing work, and practical work of ongoing work, while scheme marking was used to award marks to class tests specially prepared, practical tests specially prepared, terminal examinations and project work. Scheme marking was said to be time consuming. Commenting on the fairness of impression mark and mark scheme, both methods were thought by the teachers to be fair especially when teachers' work load is taken into consideration.

9.60 The Relationship Between Continuous Assessment Scores and Final Examination Scores.

The correlation coefficients between continuous assessment marks and the final examination marks for biology have always been positive and high in Tanzania. For example, Kitosi (1982) reported a correlation coefficient of 0.98 between the 1981 CA scores and FE scores in O-level biology; while Idama (1983) reported correlation coefficients ranging between 0.40 and 0.85 for the seven compulsory 1980 O-level subjects (Siasa, Kiswahili, English, history, geography, maths and biology). Biology had the highest correlational value of 0.85 while Kiswahili had the lowest value of 0.40.

The remaining sections of this Chapter examine in more detail the relationship between CA scores and FE score in O-level Biology using the 1983 CA and FE scores.

9.61 Performance in the Tanzanian 1983 Continuous Assessment in O-level Biology.

The raw continuous assessment scores for the same 815 candidates as earlier, and the Theory and Practical examination scores which were used in the previous chapters, constituted the input data for all analyses carried out in the present chapter.

The CA scores have been analysed through 'Sub-programme-Frequencies'; 'Sub-programme-Pearson Corr'; 'Sub-programme-Factor'; and 'Sub-programme-T-Test' of the SPSS (Nie et.al 1975); independently, and in comparison with the scores on the Theory and Practical examinations in order to examine the relationship between CA scores and Final Examination (FE) scores in the 1983 O-level biology; to make comparisons of girls' and boys' scores in CA and FE and to investigate whether 'theoretical mode of performance'; 'practical mode of performance' and 'CA mode of performance' constitute separate constructs.

9.62 Results and Discussion.

Table 9.02 shows the overall distribution of raw scores within CA and FE. FE scores, which are combined Theory and Practical examination scores, have been calculated by the formula:-

$$FE = \frac{(\text{Theory Score} + \frac{1}{2}\text{Practical Score})}{150} \times 100 \times \frac{1}{2}$$

This computation is necessary to conform with the weighting accorded to the total examination scores in the determination of final student grades in biology.

As can be seen, the CA scores are much higher, and with a wider range than the FE scores. The mean score for CA is more than double that of FE. That is, the means for CA and FE are 25.95 and 11.02 respectively.

Table 9.03 and Table 9.04 show the distribution of scores within CA and FE for girls and boys respectively. There is very little difference between the mean CA scores of 25.70 and 26.07 for girls and boys respectively. The difference being 0.37 in favour of boys. However, as with the overall distribution, the CA scores are much higher, and have a wider range than the FE scores for both girls and boys.

Table 9.02

The overall distribution of raw CA and FE Scores within the sample of 815 candidates.

Continuous Assessment

CA raw score	Absolute frequency	Cumulative percentage
6	1	.1
8	2	.4
9	5	1.0
10	4	1.5
11	6	2.2
12	7	3.1
13	6	3.8
14	13	5.4
15	13	7.0
16	12	8.5
17	20	10.9
18	30	14.6
19	25	17.7
20	40	22.6
21	32	26.5
22	42	31.7
23	42	36.8
24	43	42.1
25	45	47.6
26	38	52.3
27	48	58.2
28	32	62.1
29	41	67.1
30	40	72.0
31	37	76.6
32	38	81.2
33	32	85.2
34	30	88.8
35	31	92.6
36	15	94.5
37	15	96.3
38	12	97.8
39	7	98.7
40	3	99.0
41	5	99.6
42	1	99.8
43	1	99.9
44	1	100.0

Maximum Possible = 50.00
Mean score ($\bar{x}$) = 25.95
Standard Deviation = 6.87
range = 38.00

(N = 815).

Final Examination

FE raw scores	Absolute frequency	Cumulative percentage
3	13	1.6
4	31	5.4
5	52	11.8
6	51	18.0
7	75	26.9
8	67	35.5
9	71	44.2
10	60	51.5
11	72	60.4
12	49	66.4
13	41	71.4
14	45	76.9
15	32	80.9
16	30	84.5
17	26	87.7
18	29	91.3
19	18	93.5
20	20	96.0
21	14	97.7
22	7	98.5
23	5	99.1
24	3	99.5
25	3	99.9
28	1	100.0

Maximum Possible = 50.00
Mean score ($\bar{x}$) = 11.02
Standard Deviation = 4.84
Range = 25.61

(N = 815)

Table 9.03

The distribution of raw CA and FE scores within the Girls' sample (N = 258).

Continuous Assessment

CA Raw Score	Absolute frequency	Cumulative percentage
6	1	0.4
9	3	1.6
10	1	1.9
11	5	3.9
12	4	5.4
13	2	6.2
14	5	8.1
15	4	9.7
16	7	12.4
17	8	15.5
18	15	21.3
19	5	23.3
20	10	27.1
21	13	32.2
22	13	37.2
23	7	39.9
24	8	43.0
25	7	45.7
26	14	51.2
27	13	56.2
28	9	59.7
29	14	65.1
30	6	67.4
31	16	73.6
32	13	78.7
33	12	83.3
34	12	88.0
35	7	90.7
36	4	92.2
37	5	94.2
38	8	97.3
39	4	98.8
41	2	99.6
44	1	100.0

Maximum Possible = 50.00
Mean Score ($\bar{x}$) = 25.70
Standard deviation = 7.68
range = 38.00

(N = 258)

Final Examination

FE Raw Score	Absolute frequency	Cumulative Frequency
3	4	1.2
4	17	8.1
5	18	15.1
6	14	20.5
7	17	27.1
8	13	32.2
9	23	41.1
10	7	43.8
11	15	49.6
12	12	54.3
13	11	58.5
14	13	63.6
15	9	67.1
16	12	71.7
17	15	77.5
18	16	83.7
19	8	86.8
20	14	92.2
21	7	95.0
22	4	96.5
23	5	98.4
24	3	99.6
28	1	100.0

Maximum Possible = 50.00
Mean Score ($\bar{x}$) = 12.18
Standard Deviation = 5.76
range = 25.61

(N = 258)

Table 9.04

The Distribution of raw CA and FE Scores within the Boys' Sample (N = 557)

Continuous Assessment

CA raw Score	Absolute frequency	Cumulative frequency
8	2	0.4
9	2	0.7
10	3	1.3
11	1	1.4
12	3	2.0
13	4	2.7
14	8	4.1
15	9	5.7
16	5	6.6
17	12	8.8
18	15	11.5
19	20	15.1
20	30	20.5
21	19	23.9
22	29	29.1
23	35	35.4
24	35	41.7
25	38	48.5
26	24	52.8
27	35	59.1
28	23	63.2
29	27	68.0
30	34	74.1
31	21	77.9
32	25	82.4
33	20	86.0
34	18	89.2
35	24	93.5
36	11	95.5
37	10	97.3
38	4	98.0
39	3	98.6
40	3	99.1
41	3	99.6
42	1	99.8
43	1	100.0

Maximum Possible = 50.00
Mean score ($\bar{x}$) = 26.07
Standard deviation = 6.47
range = 35.00

(N = 557)

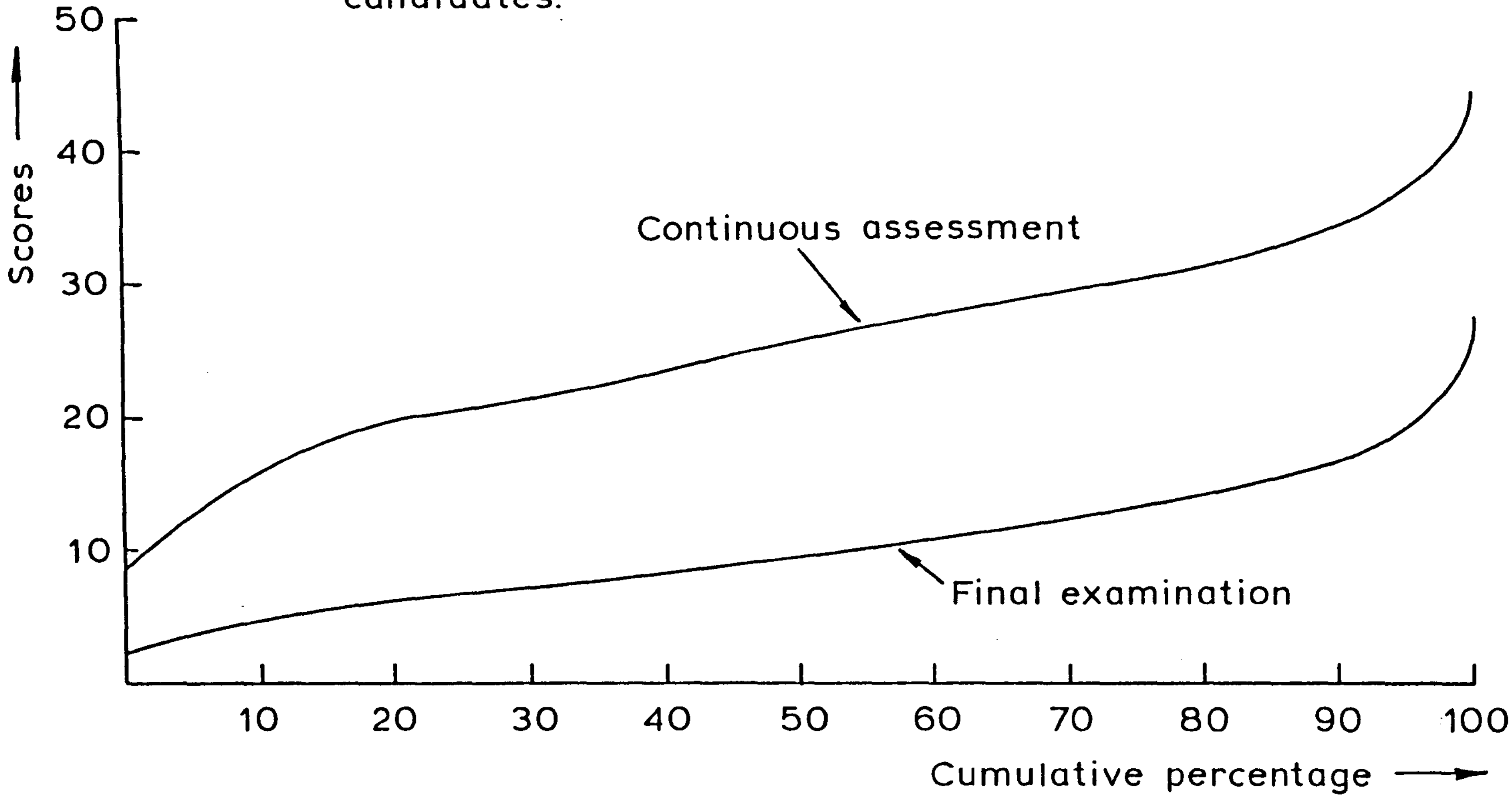
Final Examination

FE raw Score	Absolute frequency	Cumulative percentage
3	9	1.4
4	14	4.1
5	34	10.2
6	37	16.9
7	58	27.3
8	54	37.0
9	48	45.6
10	53	55.1
11	57	65.4
12	37	72.0
13	30	77.4
14	32	83.1
15	23	87.3
16	18	90.5
17	11	92.5
18	13	94.8
19	10	96.6
20	6	97.7
21	7	98.9
22	3	99.5
25	3	100.0

Maximum Possible = 50.00
Mean score ($\bar{x}$) = 10.48
Standard deviation = 4.24
range = 22.00

(N = 557)

Fig. 9.01. The relationship between score frequencies and cumulative frequencies (also called cumulative percentage) for continuous assessment and final examination for 815 candidates.



The differences between the mean scores of CA and FE for the overall sample (14.93); the girls' sample (14.52); and the boys' sample (15.59), were tested for statistical significance through 'Sub-programme- T-test' of the SPSS (Nie et.al 1975). The differences were all significant beyond the 0.001 level of significance. The difference of 0.37 between the girls' and boys' mean CA scores was not significant at the 5% level of significance. The comparison between performance in FE and CA by the overall sample, and the girls' and boys' samples is summarized in Table 9.04a and Table 9.04b while Fig. 9.01 shows a summary of the relationship between score frequencies and cumulative frequencies (also called cumulative percentage), for the CA and FE scores for the whole sample.

Table 9.04a

Comparison between performance in FE and CA by the overall sample, and the Girls and Boys Samples.

Type of Assessment	Overall Sample (N=815)		Girls' Sample (N=258)		Boys'Sample (N=557)	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard Deviation
Final Examination (FE)	11.02	4.84	12.18	5.76	10.48	4.24
Continuous Assessment (CA)	25.95	6.87	25.70	7.68	26.07	6.47

Table 9.04b

The mean sex differences and the sex favoured by the difference in performance in the FE and CA.

Type of Assessment	Mean difference	In favour of
Final Examination (FE)	1.70*	girls
Continuous Assessment (CA)	0.37	boys

* difference is significant beyond the 0.05 level of significance.

It has been suggested that greater motivation and industry on the part of the girls is a possible cause for better performance. It is often said that girls tend to work more painstakingly and steadily than boys (Hoste and Bloomfield 1975).

9.70 The Validity of CA Scores Compared with the Theory and Practical Examination Scores.

Since it was not within the scope of this study to obtain questions and items used by teachers in exercises, tests and projects for CA, it has not been possible to determine the content validity of CA. However, the construct validity has been determined through both the multitrait-multimethod matrix and factor analysis approaches described earlier.

9.71 Construct Validity by the Multitrait-multimethod Matrix Approach.

Using the multitrait-multimethod matrix approach, correlations within the different components of CA, and between CA, Theory and Practical examinations components have been calculated using 'Sub-programme Pearson Correlations' of the SPSS (Nie et.al 1975).

As explained in Chapter Eight, high intercorrelation between components is an indication that the components are measuring the same traits. On the other hand, low intercorrelation is an indication that the components are measuring different traits.

9.72 Results and Discussion.

Table 9.05 shows the intercorrelations within the different components of CA. The exercises and tests components correlate highly, and also correlate even more highly with the entire CA scores (convergent validity), with values ranging from 0.784 to 0.922. However, the projects component, consistently, shows low correlation coefficients with every other component, and with the entire CA scores, (divergent/discriminant validity), with values ranging from 0.148 to 0.340.

Table 9.05

A Multitrait-multimethod matrix for the Continuous Assessment showing the intercorrelations between the Continuous Assessment Components.

Components	Tests	Projects	Entire CA
Exercises	.784	.148	.866
Tests		.196	.922
Projects			.340

Table 9.06 shows the intercorrelations between the CA components and the Theory and Practical examinations components. The inter-correlations are of different orders, suggesting that the components are dissimilar. The correlation between the project component with every component of the Theory and Practical examinations is consistently low with values ranging from 0.023 to 0.273 only. In addition, the essay and the true-false components of the Theory examination show low correlations with every component of CA with values ranging from 0.273 to 0.356 and 0.023 to 0.262 for the essay and true-false components respectively. The inter-correlations between the experimental investigation component of the Practical examination are also low with values ranging between 0.149 and 0.488.

In order to determine the degree of divergent or convergent validity, a within and between mean intercorrelation for each component of the Theory, Practical and CA was calculated. As pointed out in Section 8.61, the within mean intercorrelation coefficient of a component is the mean of its intercorrelations with the other components of the same examination (for example, the essay with 5 components of the Theory); while the between mean inter-correlation coefficient of a component is the mean of its intercorrelations with the components of the other examinations (For example, the essay with 7 components of Practical and CA). The procedure for the calculation of the mean intercorrelation described in Section 8.61 was, again, followed to calculate the mean intercorrelations shown in Table 9.06a.

Except for the observational investigation section of the Practical paper, all the between mean components' intercorrelations are lower than the corresponding within mean components' intercorrelations. These findings seem to demonstrate divergent/discriminant validity. The following components can be said to show most divergent validity: the projects component of CA; the essay and true-false components of the Theory examination; and the experimental investigation component of the Practical examination.

The divergent validity discussed above, suggest that the components under consideration are measuring different traits of attainment. The precise nature of attainment being measured by these components is the focus of the next section where the components' scores have been factor-analysed. However, the entire CA seem to correlate highly with both the entire Theory and entire Practical examinations, with values of 0.725 and 0.630 respectively. This observation seems to give support to the teachers' use of 'practical work' and 'classroom exercises' to obtain CA marks (See Section 9.50, also Table 9.01).

Table 9.07 shows the intercorrelations within the CA components by different sexes; while Table 9.08 shows the intercorrelations between the CA components and the components of the Theory and Practical examinations for the different sexes. The pattern of the intercorrelation coefficients for both sexes is similar to that discussed above for Table 9.05 and Table 9.06. However, the main differences are the magnitudes of the intercorrelation coefficients. The girls' sample shows higher intercorrelation values than either the boys or the overall samples (Table 9.08). These intercorrelations are further discussed in Section 9.74.

The summary of the intercorrelation values between the CA scores, and the FE scores for the whole sample, girls' and boys' samples, is shown in Table 9.09. All the values, 0.753, 0.837; and 0.720, respectively, are high.

Table 9.06

A multitrait-multimethod matrix for the Final Examination (FE) and Continuous Assessment (CA) Scores showing the intercorrelations between the Theory, Practical and Continuous Assessment Components.

	Paper Components	Continuous Assessment Components: Exercises	Tests	Projects	Entire CA
Theory Examination Components	Essay	.305	.295	.273	.356
	Multiple choice	.441	.525	.216	.525
	Matching	.513	.573	.181	.585
	True-false	.212	.262	.023	.247
	free-response	.627	.691	.111	.702
	Entire theory	.639	.700	.204	.725
Practical Examination Components	Experimental Investigation	.466	.458	.149	.488
	Observational Investigation	.538	.608	.222	.629
	Entire Practical	.578	.597	.203	.630

Table 9.06a

The within and between mean intercorrelations of the Theory, Practical and Continuous assessment components, and the Entire Theory, Practical and CA components.

Paper	Paper Component	Within mean intercorrelations	Between mean intercorrelations
Theory Paper	Essay	.339	.308
	Multiple choice	.460	.411
	Matching pair	.516	.480
	True-false	.270	.180
	Free-response	.650	.572
	Entire theory Examination	.720	.600
Practical Paper	Experimental Investigation	.775	.354
	Observational Investigation	.518	.524
	Entire Practical	.846	.528
Continuous Assessment	Exercises	.685	.490
	Tests	.738	.532
	Projects	.225	.185
	Entire CA	.795	.554

The mean intercorrelations have been calculated using a Table of Transformation of r to Zr.
(McCall, 1980, p.274).

Table 9.07

Multitrait-multimethod matrices for Continuous Assessment showing the intercorrelations between the Continuous Assessment components for the different sexes.

a) Girls

Continuous Assessment Component	Tests	Project	Entire CA
Exercises	.872	.418	.866
Tests		.436	.889
Project			.596

b) Boys

Continuous Assessment Component	Tests	Project	Entire CA
Exercises	.728	.034	.868
Tests		.109	.943
Project			.237

Table 9.08

Multitrait-Multimethod matrices for the Final Examination and Continuous Assessment Scores showing the intercorrelations between the Theory, Practical and Continuous Assessment components by the different sexes.

a) Girls

	Paper Components	Continuous Assessment Components: Exercises	Tests	Project	Entire CA
Theory	Essay	.460	.484	.301	.543
	Multiple Choice	.610	.655	.291	.635
	Matching-Pair	.677	.690	.403	.707
	True-false	.252	.284	.117	.270
	Free-response	.745	.758	.327	.779
	Entire Theory	.780	.804	.386	.832
Practical	Experimental Investigation	.666	.671	.276	.660
	Observational Investigation	.719	.734	.401	.781
	Entire Practical	.736	.744	.334	.749

b) Boys

	Paper Components	Continuous Assessment Components: Exercises	Tests	Project	Entire CA
Theory	Essay	.233	.215	.262	.271
	Multiple choice	.337	.452	.203	.464
	Matching-Pair	.421	.514	.098	.520
	True-false	.186	.246	.005	.237
	Free-response	.571	.663	.029	.666
	Entire theory	.563	.648	.135	.669
Practical	Experimental Investigation	.449	.451	.004	.473
	Observational Investigation	.434	.541	.171	.550
	Entire Practical	.547	.603	.090	.624

Table 9.09

Comparison of Correlational Coefficients between Final Examination Scores for Overall performance and Girls and Boys performances.

	CA		
	Overall Correlational Coefficient	Girls Correlational Coefficient	Boys Correlational Coefficient
FE	.753	.837	.720

9.73 Construct Validity by the Factor Analysis Approach.

Tests which measure the same constructs should load highly on the same factor, while tests measuring different constructs should load highly on different factors. Therefore if CA components constitute separate constructs, three different factors should emerge, with exercises' scores, tests scores' and projects' scores loading highly, on each of the three factors. Also, if 'CA', 'Theoretical' and 'Practical' biology performance constitute separate modes of performance, upon factor analysis of the entire examination performance, three factors should emerge, with CA scores loading highly on one, Theory examination scores on another, and Practical examination scores on another.

The factor programme was run first to extract initial principal factors, their Eigenvalues and corresponding proportion of variance for the CA components only. Only one initial principal factor with Eigenvalue greater than unity was extracted for the scores of the whole sample, and for the girls and boys scores when entered into the program separately. The single factor Eigenvalue and corresponding proportion of variance for the whole sample, girls and boys samples are shown in table 9.10, and Table 9.12 respectively.

Since a rotated factor matrix could not be performed on a single factor, the factor loadings of the components of the CA scores on the single factor with Eigenvalue greater than unity are shown in Table 9.11, and Table 9.13 for the whole sample, and for the girls and boys samples respectively.

Table 9.10

Eigenvalue and corresponding variance for the single initial principal factor with Eigenvalue greater than unity extracted by the factor analysis of the CA components.

Factor	Eigenvalue	Percentage of variance	Cumulative Percentage
1	2.804	70.1	70.1

Table 9.11

Factor loading by components of the CA on the single factor extracted by the factor analysis of the CA components. Only loadings greater than 0.30 are shown.

Component	Loadings on Factor 1
Exercises	.906
Tests	.870
Projects	-
Entire CA	.999

Table 9.12

Eigenvalue and corresponding variance for the single initial principal factor with Eigenvalue greater than unity extracted by the factor analysis of the CA components for the different sexes.

GIRLS

Factor	Eigenvalue	Percentage of variance	Cumulative percentage
1	3.088	77.2	77.2

BOYS

Factor	Eigenvalue	Percentage of variance	Cumulative percentage
1	2.725	68.1	68.1

Table 9.13

Factor loading by components of the CA on the single factor extracted by the factor analysis of the CA components. Only loadings greater than 0.30 are shown.

GIRLS

Component	Loadings on Factor 1
Exercises	.895
Tests	.919
Project	.519
Entire CA	.992

BOYS

Component	Loadings on Factor 1
Exercises	.941
Tests	.850
Project	
Entire CA	.999

The factor programme was then run to extract initial principal factors, their Eigenvalues and corresponding proportion of variance for the CA, Theory and Practical examination components treated together for the whole sample, and for the girls' and boys' samples.

Three (3) initial principal factors with Eigenvalues greater than unity were extracted from the whole sample; two (2) initial principal factors from the girls' sample; and three (3) initial principal factors from the boys' sample. These factors, their Eigenvalues and corresponding proportions of variance are shown in Table 9.14, and Table 9.16 for the whole sample, and the girls' and boys' samples respectively, while the loadings on varimax rotated factor matrix for the three, two, and three factors for the whole sample, and the girls' and boys' samples are shown in Table 9.15 and Table 9.17 respectively.

9.74 Results and Discussion.

Overall the factor loadings reported in Table 9.11 and Table 9.13 for the whole sample, and for the girls' and boys' samples, for exercises, tests and entire CA scores are high with values ranging from 0.850 to 0.999. The same components (exercises, tests and entire CA), correlated highly with each other in the multitrait-multimethod matrix (Table 9.05 and Table 9.07). Bearing in mind that items which load highly on a particular factor are those which are most accurately assessing the aspect of performance represented by the factor, it can be said that both the multitrait-multimethod matrix approach and factor analysis approach to determining construct validity, point to the conclusion that the exercises and tests components of CA are not

Table 9.14

Eigenvalues and corresponding variance for the 3 initial principal factors with Eigenvalues greater than unity extracted by the factor analysis of the Theory, Practical and Continuous Assessment Components.

Factor	Eigenvalue	Percentage of variance	Cumulative Percentage
1	6.859	52.8	52.8
2	1.181	9.1	61.8
3	1.096	8.4	70.3

Table 9.15

Factor Loadings by components of the Theory, Practical and CA on varimax rotated factor matrix of the 3 principal factors extracted by the factor analysis of the Theory, Practical and CA components. (Only loading values greater than 0.30 are shown).

		Loadings on Varimax rotated factor matrix		
Paper	Component	Factor 1	Factor 2	Factor 3
Theory	Essay	.559		
	Multiple-choice	.574	.324	
	Matching pair	.625	.350	
	True-false	.368		
	Free-response	.730	.439	
	Entire Theory	.883	.371	
Practical	Exp'tal Investigation			.932
	Obs'nal Investigation	.597	.378	.300
	Entire Practical	.346	.320	.860
CA	Exercises	.300	.801	
	Tests	.377	.828	
	Project			
	Entire CA	.398	.838	

separate constructs. That is, they do not measure different traits. However, the projects component which has low correlation values with the other components and with the entire CA as a whole in the multitrait-multimethod matrix approach, does not load significantly on the single Factor extracted from the CA scores for the whole sample, and boys sample, although the project scores load significantly on the Factor extracted from the girls sample with a value of 0.519.

With regard to the CA components treated together with the Theory and

Practical examination components, the factor loadings reported in Table 9.15 and Table 9.17 range between 0.300 and 0.932 for the whole sample, from 0.386 to 0.914 for the girls' sample, and from 0.324 to 0.907 for the boys' sample.

The CA, Theory and Practical examinations components loading patterns for the whole sample (Table 9.15) have two interesting features: First, except for the experimental investigation section of the Practical examination and the project section of the continuous assessment, all other components of the Theory, Practical and CA commonly load significantly on Factor 1. Secondly, however, the Theory, Practical and CA components load on the three factors in three distinctive patterns in that, the components of the Theory examination load more highly on Factor 1, those for the Practical examination load significantly on Factor 3, while those for CA components load more highly on Factor 2.

The emergence of the three factors and the factor loading patterns of the components scores of the CA, Theory and Practical examinations on the three factors does, to some extent, suggest that 'Continuous Assessment', 'Theory' examination, and 'Practical' examination are separate modes of assessment.

Table 9.16

Eigenvalues and corresponding variance for the 2 and 3 initial principal factors with Eigenvaluesgreater than unity extracted by the factor analysis of the Theory, Practical and CA components for girls and boys respectively.

GIRLS

Factor	Eigenvalue	Percentage of variance	Cumulative percentage
1	8.144	62.6	62.6
2	1.011	7.8	70.4

BOYS

Factor	Eigenvalue	Percentage of variance	Cumulative percentage
1	6.461	49.7	49.7
2	1.368	10.5	60.2
3	1.106	8.5	68.7

Table 9.17

Factor loadings by components of the Theory, Practical and CA on varimax rotated factor matrix for the 2 and 3 factors extracted by the rotated factor analysis of the Theory, Practical and CA components for the girls and boys respectively. Only loadings greater than 0.30 are shown.

GIRLS

		Loadings on varimax rotated factor matrix	
Paper	Component	Factor 1	Factor 2
Theory	Essay	.503	
	Multiple-choice	.583	.386
	Matching pair	.607	.448
	True-false		
	Free-response	.662	.575
	Entire-theory	.767	.554
Practical	Exp'tal Investigation		.905
	Obs'nal Investigation	.624	.524
	Entire practical	.400	.914
CA	Exercises	.739	.481
	Tests	.775	.481
	Project	.509	
	Entire CA	.876	.411

BOYS

		Loadings on varimax rotated factor matrix		
Paper	Component	Factor 1	Factor 2	Factor 3
Theory	Essay	.556		
	Multiple-choice	.566		
	Matching-pair	.655		
	True-false	.355		
	Free-response	.707	.435	.302
	Entire-theory	.877	.352	
Practical	Exp'tal Investigation			.895
	Obs'nal Investigation	.626		.360
	Entire practical	.347	.318	.841
CA	Exercises		.813	
	Tests	.337	.835	
	Project			
	Entire CA	.350	.879	

When the girls' and boys' CA, Theory and Practical examinations components scores were entered into the program separately, two initial principal factors for girls, and three initial principal factors for the boys were extracted. The two, and three factors, their Eigenvalues and corresponding proportion of variance are shown in Table 9.16.

The loading patterns on the two factors for the girls' sample, and the three factors for the boys' sample are shown in Table 9.17. For the girls' scores, the Theory examination components and CA components, including the projects component, load significantly on Factor 1. The experimental investigation section of the Practical examination distinctively loads highly on Factor 2, while the observational investigation section of the Practical examination loads highly on both Factor 1 and Factor 2.

For the boys' scores, the three factor pattern which has emerged is similar to those shown in Table 9.15 for the whole sample. Again, the project's component scores do not load significantly on any of the three factors.

Although the Practical examinations' components for the girls' sample loaded on the same factor as the Theory examination components when the Theory and Practical examinations' components were factor analysed in Section 8.88 (see Table 8.36), here, with the introduction of the CA components, the Theory and Practical components load on separate factors (Table 9.17). This change in the loading pattern particularly involves the experimental investigation section which loads only on Factor 2 (See Table 9.17). A tentative explanation for this change in the loading pattern can be offered by a consideration of the intercorrelations of Theory and Practical scores with CA components. Table 9.18 shows the between mean intercorrelations of the Theory components with Practical examination components. The between mean intercorrelations of the Theory components with CA components (that is, e.g. essay with the 3 CA components and total score), are higher than the between mean intercorrelations of the Theory components with Practical components (that is, e.g. essay with the 2

Practical components and the total practical score).

It is the components with higher between mean intercorrelations, (Theory and CA components), which have loaded together on Factor 1, while the components of the Practical examination, particularly the experimental investigation section, have loaded separately on Factor 2.

Table 9.18

The between mean intercorrelations of the Theory components with 4 CA components and with 3 Practical components. (Girls' scores only.)

Theory component	Between mean intercorrelation with 4CA components	Between mean intercorrelation with 3 Practical components
Essay	.450	.305
Multiple-choice	.560	.410
Matching-pair	.630	.480
True-false	.235	.170
Free-response	.680	.580
Entire theory	.730	.600

The three factors which emerged from the boys'scores support the suggestion of the existence of the constructs of 'Theoretical', 'Practical' and 'Continuous assessment' modes of performance. It should be noted that, this is the first time it is being suggested that performance in continuous assessment could be a separate mode of performance. The projects component of continuous assessment, which has consistently produced low correlations with all other components of the examinations is further discussed in Section 10.70.

CHAPTER TEN

OVERALL DISCUSSION, INTEGRATION AND CONCLUSIONS

The previous chapters have described in detail some aspects of the role and performance of an O-level biology examination as an instrument for curriculum evaluation. The O-level biology curriculum, and 1983 O-level biology examinations in Tanzania have been the main data source for the study.

It is probably fair to say that this is the first time a Tanzanian secondary school curriculum and public examination has been subjected to such detailed evaluative analysis. Factor analysis of a public examination in biology has been reported in detail before only in Hoste's analysis of a CSE biology examination (Hoste 1977). The present chapter gives a brief overview of the findings of the present study.

10.10 Identification of topics of difficulty.

In the present study, as outlined in Sections 5.30 to 5.70, the identification of topics of difficulty in the Tanzanian O-level Biology Curriculum was done through examination item analysis for the determination of actual difficulty, and through student and teacher opinion questionnaire for the determination of perceived difficulty.

The first point to note from the items' statistics is the wide variation in the facility and discrimination indices across the 75 items of the Theory examination, and the 19 items of the Practical examination. The facility and discrimination indices for the objective items, (multiple-choice, matching-pair and true-false), of the Theory examination are within the generally acceptable ranges of 0.30 to 0.80 for F-values and 0.20 to 0.30 for D-values reported by Duckworth and Hoste (1976), Mathews (1974), Willmott and Nuttall (1972) (See Tables 5.04 to 5.06).

However, the discrimination indices for the free-response items of the Theory examination, and the items of the observational investigation section of the Practical examination are much lower than the acceptable ranges of 0.60 to 0.70 reported by Willmott and Hall (1975)(See Tables 5.07 to 5.11 and Tables 5.13 and 5.14). The source of the low discrimination indices in some of the items cannot totally be attributed to the low spread of marks, since the standard deviations for most of these items are similar.

On the other hand, the discrimination indices of the items of the experimental investigation section of the Practical examination are much higher than the acceptable values of 0.60 to 0.70 (Table 5.12). Overall, it would seem likely that the low discrimination indices of the free-response section of the Theory examination and the observational investigation section of the Practical examination is a reflection of the low standard deviations, which in turn could be caused by a group of 'less able' (in terms of total examination score), candidates responding correctly to the items, while the high discrimination indices of the items of the experimental investigation section of the Practical examination could be due to difficult items being answered correctly only by the more able candidates.

An important aspect of item analysis is the part that question choice plays in affecting the candidates' final score. Question choice was allowed in Section I of the Theory examination (See Appendix IIIa), in which the candidates were required to choose and answer only one of the three essay questions. It has been pointed out that the most significant features of the essay questions are the freedom of response allowed to examinee and the fact that not only can no single answer be listed as correct and complete and given to clerks to check, but even an expert cannot usually classify a response as categorically right or wrong. 'Rather there are different degrees of quality or merit which can be recognized' (Willmott and Hall, 1975).

The two most important characteristics of question choice identified by

Willmott and Hall are the spread of marks and level of difficulty. They identify three major sources of variation between the facility of questions:

1. The syllabus topics to which questions relate may give rise to differences between questions, since it is unlikely that all topics are mastered with the same ease.
2. Questions may differ in the level of thought required to answer them; for example, one question may require the candidate simply to recall the major points of a particular issue, whilst another question may require the candidate to analyse and contribute some 'original' thinking on a particular problem.
3. The questions may differ simply in their inherent difficulty, two questions set on the same topic and at the same level of thought may still differ because one question is an easier test of a candidate's mastery of the topic than the other. (Willmott and Hall, 1975).

They further explained that more able candidates benefit from attempting the questions with a wide spread of marks, while the less able candidates benefit from attempting the questions with a narrow spread of marks. Willmott and Hall noted that, in biology, 'the final scores of candidates could vary by as much as 10 to 15 marks, of the total mark, (5% to 7.5%), due simply to the question selection of the candidates'.

It was interesting to note that, in the present study, the most popular choice question (Essay question 1), had the lowest facility index. Willmott and Hall had observed the opposite with regard to biology examinations where the most popular questions were those with highest facility estimates (Willmott and Hall, 1975).

Overall, the results of the item analysis for both the Theory and Practical examinations as a whole were not found to be related in a systematic manner to the abilities (curriculum objectives) tested by the items in that, overall, the ability facility indices (AFI) ranged between 0.120 and 0.291, with the majority, 10 out of the 13 values ranging between 0.214 and 0.291 (Table 8.18). However, there was a relationship between item facilities and topic areas tested by the items in that, the topics on which the students performed worst, topic No.5 (Respiration in plants and animals); topic No.12 (Genetics); topic No.13 (Evolution); and topic No.15 (Parasites and diseases of man), were also among the topics perceived as being difficult (Table 5.23 and Table 8.19).

Although there was a high agreement between the students' and teachers' opinions with regard to topics of difficulty, (r = 0.8035) two obvious topics about which different views were held were topic No.1 (classification), and topic No.6 (transport systems in plants and animals). Johnstone and Mahmoud have suggested that, 'the difficulty of a topic as perceived by pupils will be a major factor in their ability and willingness to learn it'. Therefore, where there is a disparity between pupils' and teachers' views, the pupils' views are probably the more important one for the researcher and for the curriculum planner (Johnstone and Mahmoud, 1980).

Therefore, on the basis of the student responses the detailed sub-topic areas which have been identified as potential causes of learning/teaching difficulties, are in descending order of difficulty: Man and pathogens; the relationship between size, surface area and the need for transport system in large organisms; evidence of evolution - the process of change and natural selection; chromosomes, genes, mutation-structure of DNA, mitosis and meiosis; conservation of the natural resources- pollution of environment; types, structure and arrangement of muscles; Mendel's first law of inheritance (dominant and recessive genes, phenotype and genotype); the importance of genetics in plant and animal breeding; and the process of cell division, enlargement and differentiation.

However, some topics which were thought to be easy, for example topic No.2, (Cell structure and function) did not give rise to high performance in the examination; while topics which were thought to be difficult, for example topic No.6, (Transport systems in plants and animals), did give rise to high performance in the examination. On the other hand, some topics which were thought to be difficult, for example topic No.5 (Respiration in plants and animals); topic No.13 (Evolution); and topic No.15 (Parasites and diseases of man), did indeed give rise to low performance in the examination (Fig.5.05). Overall, the difficult subtopic areas fall into five main topic areas: parasites and diseases of man; genetics; evolution; respiration in plants and animals; and reproduction.

That the above listed topic areas pose problems in an O-level biology curriculum is not unique to candidates and teachers in Tanzania. Similar problems have been reported elsewhere. For example, problems of understanding the relationship between size and movement at O-level biology have been reported by Shayer (1974). Brumby (1979) has reported on misunderstandings of 'mutation' and 'theory of evolution by natural selection' held by post A-level students pursuing biological studies, while the difficulties of understanding the topic of 'evolution and heredity' by boys following Nuffield O-level biology course years 1-4 have been reported by Deadman and Kelly (1978); and Okeke (1980) has reported on 'cell enlargement and cell differentiation' as one of the areas where common misconceptions are held by O-level biology students in Nigeria.

The origins of the causes of the low performances could lie in one or more of the following areas: First, the subject matter of some topics is thought to be intrinsically difficult to learn. Longden has reported that, 'the topic of genetics in secondary biology seems to cause undue learning difficulties for some students' (Longden 1982). Furthermore, it has been reported that 'examination syllabuses contain topics such as photosynthesis; respiration, water relations, and heredity, which, regardless of the method of assessment appear to pose severe problems for pupils' (Simpson and Arnold, 1982).

Despite the clear evidence of difficulties which pupils experience with such topics, educators, teachers, and professional biologists alike, consider that they should be taught either because the topics describe key life processes in organisms or because they have a more general social and scientific importance (Deadman and Kelly, 1978).

Secondly, a topic for various reasons may be difficult to teach. The teacher may lack confidence and competence with some topics. Teachers may under emphasize certain areas of the curriculum or may fail to make coherent associations of the concepts and principles of one topic with those of other topics of the course (Mathews and Leece, 1976) and this may be demanded by the scheme of assessment.

Thirdly, some of the questions, or the marking of the answer scripts, may require different ability ranges of the candidates attempting each question. Thus, low performance can be due to the inherent difficulty of the questions; and as pointed out earlier, the F-values are susceptible to variations in the leniency or severity of the marker especially in the free-response items (subjective items).

Finally, a topic may be incompatible with students' level of intellectual development due to lack of level of attainment in areas recognized as pre-requisite knowledge to certain topics; lack of ability to integrate new knowledge into existing knowledge; and attitude to learning in general, and to particular subjects.

10.20 Reliability of the examinations.

'A reliability coefficient value in excess of 0.9 is considered to be satisfactory but, values in the range of 0.8 to 0.9 are acceptable with examinations covering a limited range of ability' (Willmott and Nuttall, 1975).

The chances of achieving satisfactory reliability are greatly increased if the marks are well spread. 'A standard deviation which is one-seventh or more of the Maximum possible mark, represents a very 'satisfactory spread of scores' ' (Nuttall, 1972).

In the present study, the effects of the small standard deviations of most of the components of the Theory examination were reflected in the reliabilities of these components which were all below the acceptable range of 0.8 to 0.9 (Table 6.01). The effects of the bigger standard deviations of the components of the Practical examination were reflected in the reliabilities of these components which were much more satisfactory (Table 6.02). However, both the entire Theory and entire Practical examinations have acceptable values of reliability especially when the scores on each item were standardized by dividing them by their standard deviations to compute the reliability coefficient (See Table 6.01 and Table 6.02 - standardized alpha).

In order to increase the reliability of each of the examination papers, it might be thought necessary to increase the reliability of the individual components of each of the two papers. This could be achieved by increasing the values of the intercorrelations between the components (Willmott and Nuttall, 1975). To be able to refine the examination components so that they intercorrelate more highly, the traits to be measured should be defined narrowly so that a small area is learnt and examined. That is, fewer objectives and less content should be covered in more detail.

Such a move would increase the reliability, but certainly at the same time would decrease the content validity of the total examination because only a narrow field would be covered from a large universe. The validity of each component should be a far more important consideration. A look at any subject syllabus issued by the Institute of Education, and examined by the Examinations Council, shows clearly the large universe of the content which is called for from candidates if they are to perform well on any examination set using such syllabuses.

However, as pointed out in Chapter Six, generally, with regard to examination reliability, the important points to bear in mind are: first, the term refers to the results of the examination and not to the examination itself. Therefore, an examination can have several reliabilities depending on the group involved and the conditions under which the examination is taken.

Secondly, examination scores can be reliable over different periods of time; different samples of questions; different markers; and different groups of respondents. The score can be consistent in any one of the above respects and yet not another. The appropriate type of consistency in a particular case should be dictated by the use to be made of the examination scores. 'We cannot speak of the reliability of an examination, but must speak of the reliability of an examination applied to a certain population' (Willmott and Nuttall, 1975). If the population is relatively homogenous in attainment, as the O-level and A-level populations are, it is difficult to achieve a very high degree of reliability.

Thirdly, reliability is a necessary but not sufficient condition for validity. Low reliability can be expected to restrict the degree of validity that is obtained, but high reliability provides no assurance of the presence of validity. That is, it is possible to have a very reliable examination which is not valid.

On the whole, therefore, upon comparison with reliabilities of various biology examinations reported in the studies reviewed in Section 6.20, the reliability of both the Theory and Practical examinations considered here, although capable of improvement, can be said to be within acceptable ranges, particularly when the homogenity of the sample studied (see Section 4.51) is taken into account.

Unlike Backhouse (1972), who reported that practical papers tended to yield lower reliabilities than theory examinations in the present study, the Practical paper components and the Practical examination as a whole, yielded higher reliabilities than the Theory paper components and the Theory paper as a whole (Tables 6.01 and Table 6.02). However, as in Hoste's study of a CSE biology examination, (Hoste, 1977), the components of the examinations containing items requiring constructed responses, (free-response section of the Theory examination and the Practical examination as a whole), were more reliable than the components of the Theory examination consisting entirely of objective items (for example, multiple-choice and true-false sections). Objective items are generally known to be more reliable than free-response items.

10.30 The importance of the objectives as course objectives and as examination objectives.

A study of the biology teachers' (the practitioners'), and panel members' (the policy makers') views of the importance of the 13 curriculum objectives as course objectives and as examination objectives made through rating and ranking of the objectives showed some differences in the opinions held by teachers and panel members. Most of the objectives were assigned all possible ratings and rankings (Tables 7.02 to 7.05). Although some teachers' priorities differed from panel members' priorities, (Tables 7.06 to 7.09), the abilities concerned with scientific method were accorded highest priority by both groups; that is, the ability to apply biological principles to real life experiences; ability to follow instructions in carrying out experimental and other biological procedures; ability to interpret data to draw conclusions and make inferences; and ability to analyse and solve problems. The ability to show knowledge and understanding of biological facts, concepts, principles and theories was accorded top priority. This should not be surprising because knowledge is the basis of reasoning and therefore is the key factor and a pre-requisite for scientific literacy and application.

In a study to identify some biology teachers' goals for Advanced level teaching, Carrick (1983) found similar priorities in their listing of objectives. Very great importance was placed on 'understanding of fundamental biological principles', 'accuracy in observation and recording'. Carrick suggested that these were perceived as realistic goals, and as such were rated highly for both the actual teaching situation and 'the ideal'. Ability to carry out investigations was also rated highly in the practical teaching situation, suggesting that this too was regarded as a 'realizable goal, of importance in preparing students for examinations' (Carrick, 1983). The aims of practical work in biology are further discussed in Section 10.51.

10.40 <u>The content validity of the examinations.</u>

In this study two aspects of validity were studied in detail. They were content validity and construct validity. The study on the content validity aimed at finding out the extent to which the Tanzania O-level biology examination measured a representative sample of the sixteen curriculum topic areas, and the thirteen curriculum objective areas prepared by summarizing those listed in the Syllabus (1976) and Teachers' Guide (1978) (See Table 7.01).

With regard to the classification of the questions according to the curriculum objectives and topic areas being tested by the questions, it was observed that: first, the results of the item analysis, in particular the facility indices, were not found to be related in a systematic manner to the abilities tested by the items, that is, the ability facility indices (AFI) were all of the same order with values ranging between 0.120 and 0.291 (Also See Section 10.20, and Table 8.18). However, there was a relationship between the facility of items and the topic areas tested by the items in that, the topics on which the students performed worst corresponded to the topics which they perceived as being most difficult (Table 5.23, and Table 8.19).

Secondly, the degree to which the examiners agreed in the classification of questions on the basis of the curriculum objectives that were being tested varied from question to question, although they unanimously agreed that all questions, in addition to testing other ability areas, tested the ability to show knowledge and understanding of biological facts, concepts, principles and theories.

Thirdly, as explained earlier, although Bloom's Taxonomy of Educational Objectives, (Bloom, 1956), has been used extensively by researchers and by examination and test designers, classification of questions into the categories of the Taxonomy has often proved to be problematic, not least when undertaken by respondents inexperienced in the detail of the Taxonomy (Willmott and Hall, 1975). Therefore, in this study, the different approach to the classification of examination questions which was employed led to the emergence of two types of classification: First, the intrinsic classification of the questions in which the O-level curriculum objectives and topic areas were used as the criteria without reference to other systems of classifications (see Appendices Va, Vb, VIa and VIb; also Table 7.11, Table 7.12, and Table 7.18); Secondly, the extrinsic classification of the questions in which the intrinsic classification was related to Bloom's Taxonomy of Educational Objectives (Bloom, 1956, also see Appendix IVe and Tables 7.15 to 7.17).

A disparity of opinion was observed when the examiners classified the 13 objectives into Bloom's Taxonomy of Educational Objectives (Table 7.13). It was not an easy task to classify the 13 objectives into the categories of Bloom's Taxonomy presumably because of the inherent difficulty in understanding Bloom's Taxonomy and the degree of subjectivity of the Taxonomy. Willmott and Hall have commented that, 'the differences noted between individual questions on the basis of their taxonomic classification may simply be a function of the subjectivity of the taxonomy and not of actual differences between the questions' (Willmott and Hall, 1975). Kelly has pointed out that, 'it is possible to find biological examples to fit Bloom's categories. It is far more difficult and in some cases impossible in any relevant sense to fit Bloom's Categories to acceptable and recognizable student objectives of biological education as have been defined' (Kelly, 1972). From the views expressed above, it would be obvious that during classification of questions, one faces a problem of deciding the boundary lines separating each of the six categories.

However, on the basis of the examiners' classification of the questions according to the ability areas tested, Table 7.11 shows that all the ability areas, (course objectives), were examined through the two examination papers, although to differing degrees. A comparison of teachers' and panel members' perceptions of the importance of the objectives as course objectives and as examination objectives with the actual examination priorities (Table 7.08 and Table 7.12), showed obvious mismatches with regard to objectives No.3 (ability to apply biological principles to real life experiences); No.11 (ability to explain and discuss); No.10 (ability to formulate hypotheses from numerical data, graphs, tables and other evidences); and No.13 (ability to use manual skills to set and operate laboratory apparatus). For the details of the mismatches see Section 7.54. Thus, the correlations between teachers' and panel members' rating of course objectives and ranking as examination objectives with the actual weighting of the objectives in the examination are low (Table 7.20).

On the other hand, the correlations between the teachers' rating as course objectives and ranking as examination objectives in each case with actual examination weighting are higher than those of the panel members with actual examination weighting. These findings can be said to indicate that teachers stress in their teaching what 'is principally designed to meet the requirements of the examination' (Duckworth and Hoste, 1976).

On the basis of the examiners' classification of the questions according to the topic areas being tested by the Theory and Practical examinations (see Appendices VIa and VIb), and the weighting based on the maximum possible score accorded to each topic area, (Table 7.18 and 7.19), it was obvious that some topic areas were over examined, particularly the topic on 'Nutrition in plants and animals', which carried 52.25% of the total marks of the entire examination. On the other hand, the topics of Movement, Growth and development, Cell structure and function, and Control and co-ordination systems in organisms were under examined.

Therefore, with regard to the content validity of the examination in terms of the objectives' coverage, on the bases of teachers' and panel members' perception of the importance of the objectives as examination objectives (Table 7.09), compared with the actual examination weighting (Table 7.12), it can be said that while some of the objectives were accorded the weights which were thought to be desirable by teachers and panel members, (objective No.1,No.5, No.6 and No.7), some were underweighted (objectives No.3 and No.10), and some overexamined (objectives No.11 and No.13). Overall by the classification of the questions into the six categories of Bloom's Taxonomy (Table 7.15), it seems that the knowledge and comprehension categories appear to have been covered more extensively than the higher cognitive categories. Over emphasis on the lower categories of Bloom's Taxonomy in biology examinations has been reported by Crossland and Amos (1965), Lewis (1967), Willmott and Hall (1975), Fairbrother (1975) and Meena (1978). Overall the content validity, in the absence of a specification matrix for both the course objectives and topic areas, and on the basis of the evidence discussed above, can be said to be satisfactory, although again capable of improvement.

One obvious way to effect improvement is through the development of an examination item specification matrix which would indicate the relative weight to be accorded to each topic area and curriculum objective. Such a matrix would be a valuable tool, particularly, for item constructors, and not least for teachers effecting continuous assessment.

10.50 The construct validity of the examinations.

Part One of Chapter Eight focused on two aspects of Practical biology. First, on what should be the aims of practical biology and how should practical work in biology be assessed, and secondly, whether 'practical' and 'theoretical' biology are separate modes of performance.

10.51 Aims of Practical Examinations in Biology in Tanzania

There was a very close agreement between teachers' and panel members' indications of the abilities which can best be examined through a practical examination. The teachers' and panel members' priority listing was (Table 8.03):

1. No.6 - Make and record observations accurately.
2. No.5 - Follow instructions in carrying out experimental and other biological procedures.
3. No.12 - Write accurate reports of experiments and other biological procedures.
4. No.8 - Draw biological diagrams accurately.
5. No.4 - Plan experiments.
6. No.7 - Interpret data to draw conclusions and make inferences.

These abilities can be interpreted as the aims of practical work in biology and of practical examinations in Tanzania. Overall the rank order correlation of teachers' and panel members' priorities was high (r = 0.8059). There was also a high level of agreement on the abilities which teachers and panel members felt should be examined by practical examinations and the abilities which were actually examined in the 1983 practical examination, (see Table 7.11). Similar priority listings of aims have been reported by Lynch and Ndyetabura (1982a), Beatty and Woolnough (1982), and Carrick (1983).

10.52 The Nature of the Items of the Theory and Practical Examinations

It was not easy to relate the Factors which emerged from the factor analysis to curriculm objectives. However, relating the Factors to curriculum content areas seemed much easier than relating them to the curriculum objectives.

Overall, unlike in Hoste's study of a CSE biology examination, (Hoste 1977), where there was more support for the suggestion that the constructs underlying performance were related to topic areas dealt with by the questions than for the suggestion that they tested single educational objectives in Bloom's Taxonomy, this study showed that the constructs underlying performance were related both to topic areas dealt with by the items, and to curriculum objectives, which could be ascribed to Bloom's Taxonomy to some extent.

From the multitrait-multimethod matrix approach and the factor analysis approach to determine if the components of each paper constituted separate constructs, and if performance in theoretical and practical biology examinations constitute separate constructs, it was shown that there was some evidence of divergent/discriminant validity, (Table 8.04, Table 8.05, Table 8.06, and Table 8.06a), between the Theory and Practical paper components, and the Theory and Practical examinations as a whole. However, upon factor analysis, the observational investigation section of the Practical examination identified more with the 'Theoretical' factor on which the Theory paper components and the Theory examination as a whole loaded highly, than with the 'Practical' factor on which the experimental investigation section and the Practical examination as a whole loaded highly (Table 8.15 and Table 8.15a). This could mean that, the candidates did not need practical skills or the practical setting to produce the correct responses to the items of the observational investigation section. Such items could have been included in the Theory paper, particularly if they were testing different content areas, otherwise they would have made little contribution to the function of the examinations in their present format.

Overall, it can be said that, both the multitrait-multimethod matrix approach, and the factor analysis approaches have supported the findings reported by Abouseif and Lee (1965), Head (1966), Kelly and Lister (1969), and Tamir (1972), that there are two modes of performance in biology examinations, a 'Theoretical' mode of performance and a 'Practical' mode of performance.

Welford et.al. of the Assessment of Performance Unit of the Department of Education and Science in England and Wales, have given three possible reasons for the differential performance in practical and written biology that they have observed. First, they say, it is not that pupils necessarily do better on practical tasks, rather that they perform differently because the tasks themselves are different. Secondly, practical performance is different from written performance perhaps because of the importance of handling the apparatus (thinking on one's feet) afforded by the practical situation. And thirdly, practical work provides an opportunity for success to be gained by those pupils who, for whatever reason, have difficulty in writing down anything coherent, or who have 'trouble in reading the question' (Welford et.al., 1985).

It is important to note that the findings of the present study, that the 'practical' mode of performance is distinguishable from the 'theoretical' mode of performance, are different from those reported in the other studies in that, in each of the cases reported above, there was a deliberate move to set the practical test/examination with the intention of determining the performance in the practical test/examination and comparing that with performance in a form of theory (written) test/examination also specifically set for the purpose. This was not the case with Hoste's study of a CSE biology examination (Hoste 1977), and the present study. Both studies analysed a public examination. The 1983 Tanzanian O-level biology examination is, as far as the writer is aware, the first public biology examination in which the existence of separate 'theoretical' and 'practical' modes of performance have been demonstrated.

With reference to Table 9.15, one more mode of performance in biology assessment has been identified, that is, the 'continuous assessment' mode of performance.

Practical assessment and continuous assessment in biology have organizational implications. The implications of carrying out practical assessment in biology are discussed below, as well as in particular the problems as experienced in Tanzania.

10.53 The organizational implications of carrying out practical assessment in biology.

The problems of carrying out practical assessment in biology are many.

1. There may be an acute shortage of apparatus and teaching staff, which can be referred to as 'school conditions and availability of apparatus'.

2. Head (1966), and Whittaker (1974) have explained that bearing in mind the problems of availability of apparatus and materials, examiners tend to choose examinations, (test areas), which can easily be carried out by each student without sharing of apparatus or materials. Automatically this leads to a limitation of areas to be assessed. A study of the topic areas assessed in the practical examinations in O-level examinations for 1980, 1981, 1982 and 1983 in Tanzania, will clearly illustrate the point being made above. For the four years consecutively, food tests have dominated the practical examination, in particular the carbohydrate food tests (reducing sugar, non-reducing sugar, and starch tests). Such practical tests can be conducted with inexpensive apparatus.

3. Other problems have been described in connection with continuous assessment of pupils' practical work. For example, the problem of observing individual procedures at critical moments for all the class when the numbers of students are large; the problem of application of the same criteria to contrasting types of practical activities; and the impossibility of production of adequate practical tests on all topic areas which might lead to the exclusion of some topic areas. Kelly and Lister (1969), have indicated three possible reasons for topics to be excluded: first, practical tests on some of the topics involve investigations requiring a much longer time than that available for a test, such as experiments from the topics of ecology, genetics and behaviour. Secondly, there may be a limited availability of expensive apparatus necessary for some specific experiments. Thirdly, there may be the uncertainty of the availability or performance of living materials on the day of the practical examination.

In Tanzania, in addition to the problems discussed above, lack of enough laboratory space has been pointed out by the biology teachers as causing major organizational problems, particularly on the day of practical examinations. On the basis of the responses received from the 49 schools (which represent 57.6% of the total government maintained secondary schools in Tanzania), up to 40 schools (82%) indicated that the available laboratory space could accommodate, on average, between 35 and 40 students on a normal teaching day, while the same laboratory space could accommodate only 18 to 25 candidates on an examination day. The practical examinations, therefore, are carried out in shifts. The number of candidates per school for the whole country for 1982 ranged from 64 to 150, with the majority of the schools presenting over 100 candidates. The magnitude of the problem is felt more in the larger schools, particularly in the day schools. Sometimes the involvement of the entire school staff can be called for to try to minimize the chance of those who did their examination in an earlier shift from communicating with those who have to do the examination in later shifts.

Five major issues seem to pose the greatest problems for teachers when preparing and carrying out practical examinations in Tanzania. The issues, given in order of magnitude of concern expressed by the teachers in this study are:-

1. Collection of specimens.
2. Availability of equipment.
3. Size of available laboratory space.
4. Lack of laboratory technician.
5. Lack of space for safe storage of specimens.

Directly related to the problem of the collection of specimens is the need for accurate identification of the specimens. Nineteen teachers, out of the 49 who returned their questionnaire, (that is 38.8%), indicated that they had failed, at least once, to identify a specimen required for the practical examination. However, when asked to explain how they got over the problem, they indicated that they usually consulted a fellow biology teacher at the same school before consulting a teacher in a neighbouring school or the Examinations Council, if need be. Identification of plant specimens was regarded as causing more problems than that of animal specimens.

Finally, although it was indicated unanimously by the teachers that the ability to use manual skills to set up and operate laboratory apparatus could best be assessed through a practical examination, the panel members explained that it has never been possible to allocate specific marks for this ability area because of the difficulty of assessing such activities. In order to do so, it would be necessary to observe each candidate during the examination. Such a task calls for several examiners. Probably the same number of examiners to candidates per shift. Therefore, candidates' ability to use manual skills to set up and operate laboratory apparatus has always been inferred from their overall performance in the practical examinations where the use of such skills is inevitable.

To overcome some of these difficulties, other schemes of assessment of practical biology, although they are not without problems, seem to be gaining popularity. For example, the scheme for teacher assessment of practical work in A-level biology for the Joint Matriculation Board (Sands and Forrest 1981); and the internal teacher assessment of candidates' practical notebooks for the University of London A-level biology examinations (Kingdon and Hartley, 1980, 1982), to name but a few.

Sands has stated that, in spite of the difficulties and problems which internal assessment pose, there is a growing opinion among teachers and those concerned with public examinations that the teacher is in the best position to assess his or her own students and that the science schemes in action are working well. Achieving comparibility between the assessment of different teachers can be coped with by moderation (Sands and Bishop, 1984).

Whittaker has listed five advantages of internal assessment:

1. The danger of chance failure or success are decreased by making assessment possible on a number of occasions.
2. A wide variety of means of assessment may be used covering a wide

variety of qualities.

3. It makes possible the development of assessment procedures which are suited to the facilities available in particular schools and which are closely related to the courses of practical work devised by these schools.
4. It requires teachers to think about assessment in terms of educational and assessment objectives.
5. It enables pupils to be assessed upon how they undertake practical work as well as upon its outcomes.
(Whittaker, 1974, p.95).

Most of all, it has been demonstrated by the experimental investigation section of the Practical examination in the present study that 'practical' mode of performance is distinctively different from the 'theoretical' mode of performance. Thus, the problems posed by assessment of practical work in biology are worth tolerating.

However, in Tanzania, since the observational investigation section of the Practical examination seemsto have more in common with the Theory examination than with the Practical examination, there is need to consider the re-organization of this part of the practical examination so that only items which require a laboratory setting can be included.

10.60 Performance by the different sexes.

The performance of boys and girls was considered separately as well as combined in the total sample because of the current international interest in different attitudes to science and achievement of the sexes.

Comparisons of performance in different ability areas, topic areas, of performance and the precise nature of the facets of attainment being measured by the Theory and Practical examination items were studied through factor analysis.

The performance by the different sexes in different ability areas tested by the Theory and Practical examinations, reported in Table 8.26, shows an overall girls' superiority. However, there was no indication of sex superiority with regard to performance in the topic areas tested. Both girls' and boys' results indicated that the topics found most difficult were parasites and diseases of man; evolution; genetics; and respiration in plants and animals (also see Section 10.10).

The performance on the different papers indicated that girls performed relatively better than boys in the Practical examination. However, boys performed better than girls in the multiple-choice, and true-false sections of the Theory examination. Better performance by boys than girls in objective type items has also been reported by Murphy (1978) and Harding (1981). Harding further indicates that structured questions appear to show the least sex bias.

As in the entire sample, an analysis of the facets of attainment being measured by the items of the Theory examination and the Practical examination in the girls' and boys' samples showed that the constructs underlying performance were related both to topic areas dealt with by the items, and to curriculum objectives. The factor patterns for the boys' sample were more similar to the overall sample than were the girls' factor patterns. However, for both girls' and boys' samples, for the Practical examination, the factor patterns resemble the overall factor patterns (see Table 8.33 and Table 8.12).

In interpreting the performance of the different sexes, Murphy (1978), has pointed out that there are areas of intellectual functioning which have 'traditionally' been thought of as displaying consistent differences between sexes, and there is a possible link between these and differences in examination performance. The proposed sex differences in ability include verbal ability, mathematical ability and visual-spatial ability. 'Certainly these skills have been observed to differ between the two sexes, although the extent of the differences and their dependence on the type of test used is far from being clear cut' (Murphy 1978)).

Sagara (1975) has further cautioned that care must be taken when interpreting studies which demonstrate sex differences in intellectual abilities because considerable overlap and inter-relatedness of knowledge and intellectual skills exists. She points out that there is normally more variation in the skills within groups of males or females than there is between them.

Since the Tanzania O-level biology examinations include items of different types, some of which are known to favour boys and others to favour girls, neither of the sexes can be said to have been generally favoured by the format of the examinations.

10.70 Continuous assessment.

In Chapter Nine, a continuous assessment scheme was defined as one which involves a systematic collection of marks or grades by the teacher over a period of time and the consolidation of the marks or grades into a final score taken into account in deciding the candidates' final grades. The most fundamental characteristic of continuous assessment, therefore, is the continuity of assessment over an extended period of time.

In Tanzania the continuous assessment scheme is implemented with the assumption that the student obtains a mark which reflects fairly his ability and also show him where he needs to improve his work before the final assessment by the end-of-course examinations.

Although in biology the teachers are obliged to formulate their schemes of work in line with the Syllabus (1976), and the Teachers' Guide (1978), each teacher has the freedom to prepare classes, laboratory programmes, and marking procedures to suit his own operative conditions. Other than the indication of when to give exercises, (after completing six forty-minutes teaching periods), and tests, (at the end of each term), there are no other terms of reference for the award of marks.

Thus, variation in the students' marks could be accounted for in two ways. The first is the 'real' variation due to difference in ability and motivation between individual students. The second is due to lack of comparability in the standards of exercises, tests and methods used by different teachers to award the marks. As was pointed out in Chapter Nine, marks are awarded both by impression mark and by using mark schemes.

From the teachers' questionnaire, the methods used to obtain a continuous assessment mark in biology, in order of frequency of use of the method were listed as:-

1. Class tests specially prepared for the purpose.
2. Class exercises based on ongoing work.
3. Practical exercises based on ongoing work.
4. Practical tests specially prepared.
5. Terminal examinations.
6. Project work.
7. Home work (assignments based on ongoing work).

Due to the great shortage of equipment, groupwork is the order of most practical exercises of ongoing work and projects, but each student writes individual reports.

However, the components of continuous assessment recognized by the Examinations Council are the exercises, tests and projects. Projects are included because they are assumed to allow pupils to demonstrate a wide variety of abilities, skills and characteristics which cannot be displayed in examination situations.

That the project component of continuous assessment is different from the other components of CA and also from the Theory and Practical examinations components was demonstrated in the multitrait-multimethod matrix, where the correlations between the project scores and the other CA components were all very low (Table 9.05). Correlations between project scores and the components of the Theory and Practical examinations components (Table 9.06 and Table 9.06a) were also low; and in the factor analysis of CA the project did not load significantly on the single factor extracted from the factor analysis

(Table 9.11).

It was not part of this study to determine the content validity of the continuous assessment components however, construct validity was determined through the multitrait-multimethod matrix approach, and by factor analysis. Through the multitrait-multimethod matrix approach, the intercorrelations between the exercises and tests component scores of CA were found to be very high, for the whole sample and for the girls and boys samples separately. This must be an indication of the overlap between the attainments measured in the two categories. On the other hand, as already mentioned, the intercorrelations displayed by the project component scores were all low.

Through the factor analysis approach it was further demonstrated that the components of CA, other than the project, that is, exercises and tests, are not separate constructs. However, when all these components were factor analysed together with the components of the Theory and Practical examinations, it was shown that Continuous assessment, Theory and Practical examinations components loaded significantly on three separate factors (Table 9.14). This could be an indication that there are three separate modes of performance in biology, the 'theoretical' mode of performance, the 'practical' mode of performance, and the 'continuous assessment' mode of performance.

However, the present study did not go into the details of identifying what was being measured by the components of CA, particularly the project component, and it is suggested that this is an area on which future research on continuous assessment in Tanzania might concentrate.

With regard to general performance in continuous assessment compared to the performance in the final examinations, the findings of the study, (Table 9.01, Table 9.02, Table 9.03 and Figure 9.01), have shown that the continuous assessment scores for the whole sample, and both for girls' and boys' samples alike, were much higher than the final examination scores. The differences between the continuous assessment mean score and final examination mean score were found to be statistically significant beyond the 0.001 level of significance for all three samples. The difference between the mean performance by girls and boys in continuous assessment was not found to be statistically significant at the 0.05 levelof significance. These findings are not unique to the present study. Higher performance in continuous assessment schemes than in corresponding written examinations has been reported by Chalmers and Stark (1968), Eggleston and Newbould (1969), and Hoste and Bloomfield (1975). Lack of evidence of different performance standards in CA by boys and girls has also been reported by Hoste and Bloomfield (1975).

Some lack of conformity between continuous assessment and final examination scores should be expected because the marks awarded independently by the teachers and the external examiners may not be based on exactly the same evidence (indeed this provides part of the rationale for conducting CA in the first place). The high correlation between the final examination scores and continuous assessment scores reported in the Tanzanian case must be an indication of a high degree of conformity between abilities tested during continuous assessment and those tested in the final examination. Maybe the wide involvement of teachers in examinations, has enabled the teachers to pick up cèrtain techniques of testing which they in turn use in carrying out continuous assessment.

The main problems which teachers face in implementing the continuous assessment scheme in biology were listed, in order of magnitude:

1. Shortage of time for preparing and marking exercises and tests.
2. Student absenteeism.
3. Entering marks on CA record forms.

Student absenteeism was indicated to be more pronounced in the day schools than in boarding schools. Absenteeism in the boarding schools was mainly found to be due to illness.

Continuous assessment and practical examinations can be regarded as inflictions to be endured because in spite of the problems of implementation of the continuous assessment scheme, of those interviewed, no teacher would wish to see continuous assessment and practical examinations disappear. Continuous assessment clearly has advantages and disadvantages. The effects of continuous assessment can be considered under three separate categories: effects on students; effects on teachers; and effects on the curriculum.

Through continuous assessment, students become informed of their continuous progress throughout the course. The fact that students are aware that the continuous assessment results affect their final assessment, makes them work hard and they can be persuaded to improve their level of attainment. This seems to suggest that, students who are assessed continuously should obtain better results than those, of equal ability, assessed only at the end of the course through a single examination (Hoste and Bloomfield, 1975).

Caution has been raised in the way marks should be allocated to students because a continuously low mark award can have a negative effect. However, Pickup and Anthony (1968) found out that an occasional low mark acted as a stimulus to normally high attainers, while an occasional high mark for low attainers gave some encouragement for more improved work.

On the part of the teachers, continuous assessment demands extra, careful preparation of tests, marking and recording of results. In the implementation of an assessment scheme for the Scottish HNC course in practical chemistry, Chalmers and Stark pointed out that teachers had to devote much more time in organizing, supervising and marking the work, in addition to preparing reports (Chalmers and Stark, 1968). Young has added that 'when a teacher has to prepare several examinations during a year, he may tend to get a bit hurried and write poorly phrased questions which, not even he could answer satisfactorily' (Young, 1970).

Regardless of the difficulties, many teachers regard continuous assessment as being extremely valuable educationally in that it enables due weight to be attached to some aspects of the course which cannot be assessed by written examinations (Cohen and Deale, 1977).

The potential use of continuous assessment for curriculum evaluation can best be summarized by the words of Eggleston and Kerr, who, when talking of continuous assessment in Nuffield sciences, have pointed out that 'the refinement of assessment procedures might play a wider role in curriculum development than merely as a means of measuring a pupil's attainment at the end of a course' (Eggleston and Kerr, 1969). They further say, 'provided the assessments arise directly from an adequate specification of the educational objectives, the continuous evaluation of attainment has important consequences. These include more emphasis on immediate and intermediate gains rather than on terminal or more remote outcomes; higher levels of student motivation; and encouragement to use a wider variety of teaching methods' (Eggleston and Kerr, 1969).

10.80 The Appropriateness of the Examinations as a Tool for Curriculum Evaluation.

As was mentioned in Chapter One and Two, curriculum development in Tanzania is centralized and is seen as an intergral part of the policy for educational reform - The Education for Self-Reliance (ESR).

The problems of evaluating a centralized curriculum have been discussed in Section 2.52. Performance on public examinations is one of the best indicators of the success or failure of a public curriculum, and hence analysis of factors which contribute to such success or failure is potentially a powerful evaluative tool.

In Section 3.30, three basic conditions were identified which had to be fulfilled in order that examinations might be used to evaluate the curriculum. These were that:

1. The specification of the purpose; the content; and the processes of the curriculum must be made explicitly and quantitatively in the curriculum itself.

2. The assessment system, should match the purpose, content and processes of the curriculum both qualitatively and quantitatively.

3. A substantial part of the curriculum must be studied and assessed.

Basically, for O-level biology in Tanzania, condition No.1 above has been met through what is documented and specified in the O-level biology syllabus (1976) and the Teachers' guide (1978) (See Appendices I and II). Condition No.3 can be assumed to have been met since the course is a four year one and in that the final examinations are taken at the end of the fourth year, it can be assumed that a substantial part of the curriculum had been studied.

The evaluation of the achievement of condition No.2 is what has constituted the bulk of the present study. That is, the evaluation of the appropriateness of the examinations as a tool for curriculum evaluation. Before making judgement on the appropriateness of the examinations in this context, the reader is reminded that, the O-level examinations in Tanzania 'serve a dual purpose, they are both terminal and promotional' (Education, Ministry of, 1980 p.17). Therefore, since they are terminal, the amount of knowledge and abilities acquired by the candidates in the course of the four years of study should be of significant importance to them for reasons of future citizenhood; and in terms of the selective function of the examinations it must be possible to discriminate between different levels of performance with reliability and validity, for selection of candidates for entry into the few places available in higher educational institutions and into specified careers.

So the question is, to what extent are the examinations fulfilling the two functions of evaluating the effectiveness of the course and measuring attainment in a way suitable for selecting candidates for higher education? In order to answer these questions the qualities which an examination should display for evaluation of the effectiveness of course and for selection should be differentiated.

The basic qualities of an examination in measuring effectiveness of performance and in allowing for selection, are manifested in the facility and discrimination indices of the items. A course of instruction is devised in order to achieve specified objectives. To demonstrate the relative effectiveness of instructional procedures in achieving specified objectives, a display of competence in the skills demanded by the items is required of candidates across the different content areas and ability areas of the curriculum. Therefore, with regard to suitability of examinations for evaluation of the effectiveness of a course, items which result in a spread of scores across the normal distribution curve might not be suitable because it should not be impossible, sometimes, to have perfect or almost perfect scores on some items by all pupils.

Thus, Eggleston (1974) commenting on measuring attainment for curriculum evaluation stated:

> 'If we assume that the instruction has been even partially effective the pupils with the highest scores in the post-instruction test will be among those who have achieved the objective. The scores on items which contribute to the measure of achievement of the objective will: a) have higher facility values in post-instruction tests; b) correlate positively with score totals' (Eggleston, 1974, p.232).

However, for selection purposes, differential performance by candidates based on total score is regarded as desirable, hence items with high D-values, leading to the spread of candidate scores are regarded as desirable.

If we now return to the data resulting from the analysis of the Tanzanian O-level biology examinations, and with the view of appropriateness of the examinations as a tool for the evaluation of the effectiveness of the course and for selection of candidates, it is clear that most of the mean scores were relatively low compared to the maximum possible scores for each of the items, particularly the items of the free-response section of the Theory examination (Table 5.07 to 5.11). Likewise, the item facility indices, (the ratio of the mean score to the maximum possible score), were low (Tables 5.03 to 5.14), while the discrimination indices, in addition to being high, in some cases correlated negatively with score totals (see Tables 5.06 and 5.11). The cause of the low performances could lie in any one or all of four possible origins: the curriculum; teachers; students and the examinations.

As seen from the Syllabus (1976) and Teachers' Guide (1978), (See Appendices I and II), which constitute the chief curriculum materials, the course content does not include particularly unusual content areas, compared to O-level biology curriculum in UK. The topics of ecology, evolution and genetics which have been generally shown to cause undue learning/teaching difficulties at various stages of secondary and even post-secondary levels of education (Shayer, 1974; Deadman and Kelly, 1978; Brumby, 1979; Longden, 1982); are included in the school curriculum despite the difficulties experienced, because they are thought to describe key concepts and life processes.

From the point of view of topic facility indices (TFI), it is evident from Table 5.23 that both teachers' and students' perceptions of topic facilities were high, and yet the actual student performances in the examinations on the same topic areas were low.

It was not part of this study to determine teachers' competence and effectiveness. However, some elements of teacher effectiveness, for example completion of teaching the course content, can be deduced from the responses to the Pupil Opinion Questionnaire, and Section II of the Teachers' Opinion Questionnaire, where overall coverage of the curriculum can be said to have been achieved since there was no topic area, out of the 16 topic areas, which was indicated as not having been covered fully.

Therefore, the most likely explanation for the low performances lies within the examinations. The low facility indices are indications of inherent difficulty of the items of the examinations. For example, twenty six (26) items, out of forty two (42) items of the free-response section of the Theory examination have facility indices of less than the generally accepted facility index of 0.20 for free-response items. The F-values of these 26 items range between 0.002 and 0.190 (Tables 5.07 to 5.11).

That the items required different ability ranges of candidates attempting each item is evident from the examiners' classification of the items into the different ability areas, (See Appendices Va and Vb). The ability facility indices (AFI) are low ranging from 0.098 to 0.604 with a mean ability facility index of 0.213 (Table 8.26). As already mentioned in Section 8.86, there was no significant differences between the differences in the performances of the different ability areas. So the question is, are these low facility indices (item facility indices, topic facility indices and ability facility indices), indications that the examinations did not sample the knowledge and abilities which teachers thought they had developed in their students and students thought they had acquired?

In order to make a valid statement about the performance standard, a statement of specified performance standard would seem necessary. That is, some criterion referencing. If such a statement was available, then the above results would have been easily interpreted in terms of the specified performance standards. However, although a statement of performance standard is not available, from the evidence given from the data analysis, particularly the facility and discrimination indices, it can be said that the examinations seem to have been more suited for the selection of the candidates for entry into the limited places in higher educational institutions and for selected careers placement, than for evaluating the effectiveness of the course.

These are the author's personal views; and it is worth reminding readers that Kelly in one of the evaluation studies of the Nuffield A-level biology trials, 'Overall Achievement of Students', stated, 'judgement about standards must inevitably rest with individuals. One can but provide data and ask it to be scrutinized in relation to the objectives and content of the scheme and from the methods of assessment used' (Kelly, 1971).

10.90 Recommendations and Suggestions for Further Research.

Taba has said that the translating of evaluation data into curriculum improvement or improvement of teaching is too often done rather mechanically. Low performance by students is taken as an automatic sign for a need to spend more time on the area of low performance without sufficient analysis of the possible causes for the low performance or its meaning in the total pattern of achievement (Taba, 1962). It is therefore suggested that, in addition to taking immediate action to solve some of the problems which have been pin pointed in the course of this study, further research be conducted. The following suggestions are offered:

1. For the purpose of the improvement of the curriculum and teaching/ learning situations, having isolated the main topic areas of perceived and actual difficulties, the next move should be to isolate the reasons for the existence of the difficulties. It is by so doing that methods of dealing with such difficulties can be identified and their implications on both teaching and learning be discussed effectively.
2. For the improvement of the examinations, and examination procedures, the development of an examination item specification matrix which would indicate the relative weights to be accorded to each topic area and curriculum objective would be an invaluable tool, particularly, for item constructors, and without doubt for teachers effecting continuous assessment.
3. Since the observational investigation section of the Practical examination seems to have more in common with the Theory examination than the Practical examination, there is need to re-consider and re-organize the format of the Practical examination so that only aspects which require a laboratory setting would be included. This might decrease and/or make worthwhile some of the organizational problems which teachers go through in preparing and carrying out the Practical examinations.

4. It was demonstrated that the Project component of CA is different from the other components of CA and also from the Theory and Practical examinations components. However, the details of the attainment being measured were not dealt with. Therefore, it is suggested that this is an area where future research on continuous assessment in Tanzania might concentrate.
5. Although in biology choice is offered only in the three essay questions, each of those questions carries 20% of the total weight of the Theory examination. Therefore, a well answered or poorly answered question could make considerable differences in candidates final scores. It is therefore suggested that, more research be conducted to determine the full effects of variations in the spread of marks across questions, where choice is offered, and to find out possible procedures for standardizing scores to counter any inequalities arising from such variations.
6. The results and discussions presented in this study have been based on the examination data from 815 candidates, which represents approximately 10% of the total candidate population for the 1983 O-level biology candidacy. Whether the results discussed are a reflection of a National trend in the past needs to be examined through further research and analysis of previous years' results. Also a comparison of the performances of boys and girls in other science subjects and also in the social sciences would make a useful contribution towards existing knowledge of differential performance by the two sexes, where little has been reported from Africa generally and particularly from Tanzania.

Finally it is hoped that, the research approach used to study this familiar subject in this study, can be extended to study other subjects, in Tanzania and elsewhere, to investigate the role that examinations can play in the process of curriculum development and evaluation. Not all curricula and examinations can be monitored by exactly the same processes, but it is hoped that the principles and some of the processes could be applied to the science subjects at least, but also possibly to the social science (arts) subjects.

Bibliography

Abouseif, A. A. and Lee, D. M. (1965).'The Evaluation of Certain Science Practical Tests at the Secondary School level', British Journal of Educational Psychology (35) pp.41-49.

Adams, R. M. and Murphy, R. J. L. (1982). The Achieved Weights of Examination Components, Educational Studies. Vol.8, No.1, pp.15-22.

Agazzi, A. (1967). The Educational Aspects of Examinations: Education in Europe Series II, General and Technical Education No.10. Stransbourg, pp. 65-80.

Amos, R. (1967). 'Teachers' opinions about the place of Scientific method in biology courses leading to the General Certificate of Education examination at Ordinary Level'. M.Ed.Thesis, Manchester University.

Amos, R. (1970). 'What is the place of Scientific Method in O-level Biology Course?'...some notes gained from a survey of teacher opinion, Journal of Biological Education (4) pp. 87-95.

Backhouse, J. K. (1972). 'Reliability of GCE Examinations: A Theoretical and Empirical Approach: In Nuttall and Willmott (eds), British Examinations: Techniques of Analysis. National Foundation for Educational Research. pp.89-117.

Backhouse, J. K. (1974). Pass or Fail? A study of Pass Rates at the Ordinary Level of the General Certificate of Education. NFER Publishing Company Ltd. (1974).

Beatty, J. W. and Woolnough, B. E. (1982). 'Practical Work in 11-13 Science: the context, type and aims of current practice', British Educational Research Journal Vol.8,No.1. pp.

Becher, T. and Maclure, S. (1978). The Politics of Curriculum Change. Hutchinson of London.

Bell, B. and Brook, A. (1984). Aspects of Secondary Student's understanding of Plant Nutrition: Full report. Centre for Studies in Science and Mathematics Education. The University of Leeds.

Bloom, B. S. (1956). Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook : Cognitive Domain. Longmans Green and Co.

Bloomfield, B. A; Dobby, J. L. and Kendall, L. (1979). Ability and Examinations at 16+. Macmillan Education.

Brereton, J. L. (1944). 'The case for Examinations'. Cambridge University Press.

Brumby, M. (1979). 'Problems in learning the concept of natural selection.' Journal of Biological Education. 13 (2) pp.119-122.

Butcher, H. J. (Ed.) (1968). Educational Research in Britain: London University Press. pp.167-184.

Catell, R. B. and Warburton, F. W. (1967). Objective personality and motivation tests. Urbana: University of Illinois Press.

Carrick, T. (1983). 'Some biology teachers' goals for Advanced Level teaching'. Journal of Biological Education, 17 (3) pp.205-214.

Chalmers, R. A. and Stark, J. (1968). 'Continuous assessment of practical work in the Scottish HNC course in chemistry'. Education in Chemistry, 5 pp. 154-155.

Cohen, L. and Deale, R. N. (1977). Assessment by teachers in examinations at 16+. Evans/Methuen Educational.

Comber, L. C. and Keeves, J. P. (1973). Science Education in Nineteen Countries. John Wiley and Sons, New York.

Connaughton, M. I. (1969). 'The Validity of Examinations at 16+'. Educational Research. (11) pp.163-178.

Connaughton, M. I. and Skurnik, L. S. (1969). 'The Comparative effectiveness of several short-cut item analysis procedures'. British Journal of Psychology, 39 pp. 225-232.

Cronbach, L. J. and Meehl, P. E. (1955). 'Construct validity in psychological tests'. Psychological Bulletin, 52 pp. 281-302.

Cronbach, L. J. (1963). 'Course Improvement Through Evaluation'. Teachers' College Record. No.64 pp.672-683.

Cronbach, L. J. (1970). Essentials of Psychological Testing. (3rd Ed.). Harper and Row Publishers.

Crossland, R. W. and Amos, R. (1965). 'Towards new O-level examinations in Biology'. Biology and Human Affairs. 30 (3) pp.35-39.

Deadman, J. A. and Kelly, P. J. (1978). 'What do Secondary School boys understand about evolution and heredity before they are taught these topics?' Journal of Biological Education. 12 (1) pp.7-15.

Drake, C. A. (1941). 'A short method for analysing Examination questions': School and Society'. p.392.

Duckworth, D. and Hoste, R. (1976). Question banking: an Approach through biology. School Council Examinations Bulletin 35. Evans/ Methuen Education.

Dunn, T. F. and Goldstein, L. G. (1959). 'Test difficulty, validity and reliability as function of selected multiple choice item construction principles'. Educational and Psychological measurement. (19) pp.171-179.

Ebel, R. L. (1956). 'Obtaining and Reporting Evidence on Content Validity'. Educational and Psychological Measurement, 16 pp.269-282.

Ebel, R. L. (1972). Essentials of Educational Measurement. Prentice-Hall, Inc. Englewood Cliffs. New Jersey (2nd ed.).

Education, Act of the Institute of,. (1975) Act of Parliament for the Establishment of the Institute of Education No.13 of 1975. Government Printers, Dar-es-Salaam.

Education, Institute of,. (1967b). The Biology Panel Report. University College Dar-es-Salaam. Dar-es-Salaam (mimeo).

Education, Institute of,. (1967). Recommendations of the Combined Science Panels. University College Dar-es-Salaam. Dar-es-Salaam.

Education, Institute of,. (1967, 1968, 1969). SSP Biology-Pupils' Book 1,2, 3 and 4. Dar-es-Salaam.

Education, Institute of,. (1967, 1968, 1969). SSP Biology-Teachers' Book 1, 2, 3 and 4. Dar-es-Salaam.

Education, Institute of,. (1975). History for East and Central Africa. Tanzania Publishing House, Dar-es-Salaam.

Education, Institute of,. (1978). Teachers' Guide on 'O-level Biology Syllabus. Dar-es-Salaam.

Education, Institute of,. (1982). Handbook 1982, Institute of Education. Dar-es-Salaam.

Education, Ministry of,. (1973). Heads of Secondary Schools Annual Conference 1972-1973. (Mimeo).

Education, Ministry of,. (1976). Science Syllabuses 1976 Vol.2. Printpak Dar-es-Salaam pp.52-66.

Education, Ministry of,. (1980). Basic Facts About Education in Tanzania. Printpak Dar-es-Salaam.

Eggleston, J. F. (1974). 'Measuring Attainment for Curriculum Evaluation': In Techniques and Problems of Assessment: Macintosh (ed.). Edward Arnold (Publishers) Ltd. London.

Eggleston, J. F. and Kerr, J. F. (Eds.) (1969). Studies in Assessment. English University Press.

Eggleston, J. F. and Newbould, C. A. (1969). 'Assessing Attainment in Science at CSE level'. In Studies in Assessment. Eggleston and Kerr (Eds.). English University Press. pp.16-57.

England and Wales (1944). Education Act, 1944 - Section 23.

Examinations Council of Tanzania. (1973). Act of Parliament No.21 of 1973. Government Printers. Dar-es-Salaam.

Examinations Council of Tanzania. (1981). 'Miaka Kumi ya Mitihani ya Taifa' (Ten Years of National Examinations) 1971-1981. Dar-es-Salaam.

Examinations Council of Tanzania. (1983). Examinations Statistics for November, 1981. Dar-es-Salaam.

Examinations Council of Tanzania. O-level Biology Examinations Papers 1980, 1981, 1982. Dar-es-Salaam (1980, 1981, 1982).

Examinations Council of Tanzania. O-level Biology Examinations Papers 1983. Dar-es-Salaam (1983).

Fairbrother, R. W. (1975). 'The Reliability of Teachers' Judgement of the Abilities Being Tested by Multiple Choice Items'. Educational Research. 17 (3) pp. 202-210.

Gagne, R. M. (1967). 'Curriculum Research and the Promotion of Learning'. In Perspectives of Curriculum Evaluation Tyler et.al (ed.): AERA Monograph Series on Curriculum Evaluation No.1. Rand McNally. Chicago, pp.19-38.

Gardner, P. L. (1975). Attitudes to Science: A Review. Studies in Science Education, (2) pp.1-41.

Gaskell, D. C. (1973). Curriculum development: Levels of conceptual difficulty. School Science Review, 55 (191) pp.374-378.

Gordon, P. (Ed.). (1981). The Study of the Curriculum. Batsford Academic and Educational Ltd. London.

Gronlund, N. E. (1971). Measurement and Evaluation in Teaching (2nd Ed.). The Macmillan Company. N.York.

Guilford, J. P. and Fruchter, B. (1978). Fundamental Statistics in Psychology and Education. McGraw-Hill Book Company.

Harding, J. (1981). Sex differences in Science examinations. In Kelly, A. (Ed.). The Missing Half-Girls and Science Education. Manchester University Press.

Harris, C. W. (1960). 'Tests and Examinations'. In Encyclopedia of Educational Research (3rd Ed.). The Macmillan Company, N.York.pp.1502-1517.

Hartog, P. and Rhodes, E. S. (1935). An Examination of Examinations. International Institute Examinations Enquiry. London: Macmillan.

Hartog, P. and Rhodes, E. C. (1936). 'The Marks of Examiners'. International Institute Examinations Enquiry. London: Macmillan.

Hartog, P. (1938). 'The Purposes of Examinations'. The Yearbook of Education. Evans Brothers Ltd. London. (p.5).

Head, J. J. (1966). 'An objectively marked practical Examination in O-level Biology'. School Science Review. 48 (164) pp.85-95.

Hoste, R. (1977). An Evaluation of a Certificate of Secondary Education examination in Biology. PhD Thesis. Chelsea College, University of London.

Hoste, R. (1977). 'The construct validity of some CSE Physics examinations'. School Science Review, 59 (206) pp.127-131.

Hoste, R. (1980). 'The construct validity of some CSE biology examinations: The evidence from multitrait-multimethod matrices'. Journal of Biological Education. 14 (1) pp.41-48.

Hoste, R. (1982a). 'What do examination items test? An investigation of construct validity in a biology examination'. Journal of Biological Education, 16 (1) pp.51-58.

Hoste, R. (1982b). 'The Construct Validity of some Certificate of Secondary Education Biology Examinations: The Evidence from factor analysis'. British Educational Research Journal (1982b) Vol.8, No.1, pp.31-43.

Hoste, R. and Bloomfield, b. (1975). Continuous Assessment in the CSE: Opinion and Practice. Evans/Methuen Educational.

House, E. R. (ed.) (1973). School Evaluation: The Politics and Process. McCutchan, p.3.

Howe, A. (1974). 'Formal operational thought and the high school science curriculum'. Paper presented at the 47th Annual Meeting of the National Association for Research in Science Teaching.

Hull, C. H. and Nie, N. H. (1981). SPSS Update 7-9: New Procedures and Facilities for Release 7-9. McGraw-Hill Book Company.

Idama, N. M. (1983). Relationship Between Continuous Assessment and Final Examination Scores: A case of Secondary Education Assessment. Paper presented to the Inspectors of Schools Annual Meeting. Examinations Council Mbeya.

Jenkins, D. (1976). Curriculum Evaluation: The Open University Educational Studies. Open University Press Walton Hall, Milton Keynes.

Johnstone, A. H. and Mahmoud, N. A. (1980). 'Isolating topics of high perceived difficulty in School Biology.' Journal of Biological Education. 14 (2) pp.163-166.

Johnstone, A. H. and Mahmoud, N. A. (1980). 'Pupils' problems with water potential'. Journal of Biological Education. 14 (4) pp.325-328.

Johnstone, A. H. and Mahmoud, N. A. (1981). 'Pupils' response to a model for water transport'. Journal of Biological Education, 15 (3) pp.203-208.

Jungwirth, E. (1971). 'The Pupil- The Teacher- and the Teacher's Image (some Second Thoughts on BSCS- Biology in Israel)'. Journal of Biological Education, 5(6) pp.165-171.

Kelly, A. (Ed.). (1981). The Missing Half - Girls and Science Education. Manchester University Press.

Kelly, A. (1981).'Sex differences in science achievement: some results and hypotheses'. In Kelly, A. (Ed.). The Missing Half-Girls and Science Education. Manchester University Press, pp.22-41.

Kelly, A. V. (1982). The Curriculum: Theory and Practice. Harper and Row, Publishers. London (2nd Ed.).

Kelly, P. J. (1970b).'Implications of Nuffield A-level Biological Science'. School Science Review, 52 (179) pp.272-285.

Kelly, P. J. (1971a). Evaluation Studies of Nuffield A-level Biology Trials: 1. Overall Achievement of Students'. Journal of Biological Education, 5(6) pp.315-327.

Kelly, P. J. (1971b). 'A reappraisal of Examinations'. Journal of Curriculum Studies. Vol.3, No.1 and 2, pp.119-127.

Kelly, P. J. (1971). Curriculum development and Curriculum mechanism: an account of the Nuffield A-level Biology Project in relation to the concept of Curriculum mechanism. Ph.D. Thesis. Chelsea.

Kelly, P. J. (1972). Evaluation studies of Nuffield A-level Biological Science: 2. 'Evaluation of Specific Objectives: Journal of Biological Education, (6) pp.29-40.

Kelly, P. J. and Dowdeswell, W. H. (1970). 'The Nuffield A-level Biological Science Project'. Journal of Biological Education. (4) 4 pp.251-260.

Kelly, P. J. and Lister, R. E. (1969). 'Assessing Practical ability in Nuffield A-level Biology'. In Eggleston and Kerr (Eds.). Studies in Assessment. English University Press. pp.121-152.

Kelly, P. J. and Monger, G. (1974). 'An evaluation of the Nuffield O-level biology course materials and their use: Part I'. School Science Review, 55 (192) pp.470-482.

Kerr, J. F. (1968). 'The Problems of Curriculum Reform': In Changing the Curriculum, Kerr (ed.). University of London Press.pp.13-38.

Kingdon, J. M. and Hartley, D. J. (1980).'Teacher assessment of A-level biology practical notebooks - the development of a system of moderation'. Journal of Biological Education, 14 (3) pp.245-255.

Kingdon, J. M. and Hartley, D. J. (1982). 'Teacher assessment of University of London A-level biology practical notebooks - a report on the first operational examination'. Journal of Biological Education, 16(4) pp.286-292.

Kitosi, J. A. D. (1982). The extent to which Continuous Assessment and the National Form Four Examinations are valid Measures of attainment in Biology. MA Thesis, University of Dar-es-Salaam.

Krathwohl, D. R. et al (1964). Taxonomy of Educational Objectives: the classification of educational goals. Handbook 2: Affective Domain. David McKay Co. New York.

Lawton, D. (1978). Curriculum Change in the nineteenth and twentieth centuries. Hodder and Stroughton. London.

Leece, J. R. and Mathews, J. C. (1976). 'Nuffield Advanced Chemistry Research Project'. School Science Review. 57 (198) pp.148-155.

Lennon, R. (1956). 'Assumptions underlying the use of content validity'. Educational and Psychological Measurement. 16 pp.294-304.

Lewis, D. G. (1967). 'Ability in Science at Ordinary level of the General Certificate of Education' British Journal of Educational Psychology 37 pp.361-370.

Lewis, D. G. (1974). 'Examinations' In Assessment in Education. University of London Press Ltd. pp.91-114.

Lister, R. E. (1969). 'Aims of questions in A-level biology Examinations. School Science Review, 50 (72) pp.514-527.

Longden, B. (1982). 'Genetics - are there inherent Learning difficulties?' Journal of Biological Education, 16 (2) pp.135-140.

Lucas, A. M. (1980). 'The development of a curriculum monopoly in Australian Secondary Schools: Biological Science: The Web of Life'. 1. Origin and Spread. Journal of Biological Education, 14 (1) pp.15-28.

Lynch, P. P. and Ndyetabura, V. L. (1982a). 'The Purpose of Practical Work: An Examination of the Aims of Teachers and the Perceived Influence of Practical Work According to Students in Tasmanian High Schools and Matriculation Colleges'. The Australian Science Teachers Journal, pp.80-86.

Maccoby, E. E. and Jacklin, C. N. (1975). The Psychology of Sex Differences. London, Oxford University Press.

Macintosh, H. G. (ed.) (1974). Techniques and Problems of Assessment. Edward and Arnold Publishers) Ltd. London.

Macintosh, H. G. (1976). 'Techniques of Assessment I'. In Assessment and the Secondary School teacher: Macintosh and Hale (eds.). Routledge and Kegan Paul, London, pp.25-52.

Macintosh, H. G. and Hale, D. E. (1976). Assessment and the secondary school teacher. Routledge and Kegan Paul, London.

Macintosh, H. G. and Morrison, R. B. (1969). Objective testing, University of London Press.

Maling-Keepes, J. (1978). Educational Evaluation: Key Characteristics. Australian Council for Education Research. Research Series No. 102. Hawthorn, Vic.

Mathews, J. C. (1974). 'Structured Questions' In Macintosh (ed.): Techniques and Problems of Assessment. Edward and Arnold (Publishers) Ltd.

Mathews, J. C. and Leece, J. R. (1976). Examinations, their use in curriculum evaluation and development. Schools Council Examinations Bulletin 33.

McCall, R. B. (1980). Fundamental Statistics for Psychology (3rd ed.). Harcourt Brace Jovanovich, Inc.

Meena, A. S. (Ed.). (1971). School Science Project in Biology - What is it?. Institute of Education; University of Dar-es-Salaam.

Meena, A. S. (1978). A Survey of present trends in the teaching and learning of Biology in Tanzanian Secondary Schools with proposals for improvement. PhD Thesis. University of Dar-es-Salaam.

Millman, J. and Setijadi, (1966). 'A comparison of the Performance of American and Indonesian Students on three types of test items'. Journal of Educational Research 59 pp.273-275.

Morris, N. (1961). 'An Historian's view of examinations'. In Wiseman, s (ed.): Examinations and English Education. Manchester University Press pp. 1-26.

Morrison, R. B. (1972). 'A method for analysing Choice Type Question Papers. In Nuttall and Willmott (eds). British Examinations: Techniques of Analysis : Windsor NFER pp.61-72.

Mtani, S. K. (1983). Conceptual and Practice Dimensions of Continuous Assessment: The National Examinations Council of Tanzania. Paper presented to the Sub-Regional Meeting of Institutions Responsible for Educational Assessment in East and Southern Africa. Dar-es-Salaam, (September, 1983).

Murphy, R. J. L. (1978). 'Sex differences in examination performance: do these reflect differences in ability or sex-role stereotype?'. Educational Review. 30 (3) pp.259-263.

Murphy, R. J. L. (1980).'Sex Differences in GCE Examinations Entry Statistics and Success Rates.' Educational Studies, 6 (2) pp.169-178.

Ndyetabura, V. L. and Lynch, P. P. (1982b). 'Assessment of Practical Work in Schools: An Examination of the Preferences of Teachers and Pupils in Tasmanian Schools'. The Australian Teachers Journal. pp. 78-84.

Nie, N. H; Bent, D. H; and Hull, C. H. (1975). Statistical Package for the Social Sciences (SPSS) London: McGraw-Hill (2nd Ed.).

Njabili, A. F. (1978). Problems Pertaining to the Teaching of Biology, using the present syllabus. Institute of Education. Dar-es-Salaam.

Nunnally, J. C. (1972). Educational Measurement and Evaluation. McGraw-Hill, Inc. (2nd Ed.).

Nuttall, D. L. (1972). 'The CSE Investigations'. In British Examinations: Techniques of Analysis; Nuttall and Willmott (Eds.). NFER book series. pp. 48-60.

Nuttall, D. L. and Skurnik, L. S. (1969). Examinations and item analysis Manual. Slough: NFER.

Nuttall, D. L. and Willmott, A. S. (Eds.). (1972). British Examinations: Techniques of Analysis: Windsor. NFER.

Nyerere, J. K. (1967). Freedom and Socialism. Oxford University Press. Dar-es-Salaam.

Nyerere, J. K. (1967). 'Education for Self-Reliance'. In Freedom and Socialism. Oxford University Press, Dar-es-Salaam. pp.267-290.

Okeke, E. A. C. and Wood-Robinson, C. (1980). A Study of Nigerian Pupils' Understanding of selected biological concepts. Journal of Biological Education, 14 (4) pp. 329-338.

Oliver, A. I. (1965). Curriculum Development. Dodd Mead and Co.

Ormerod, M. B. (1975). Science education, attitudes, subject preference and choice. Paper presented at the Girls and Science Education Conference, Chelsea College.

Ormerod, M. B. and Duckworth, D. (1975). Pupils' Attitudes to Science: NFER Publishing Company.

Owen, J. G. (1973). The Management of Curriculum Development. Cambridge University Press.

Parsons, C. J. (1973). Theses and Project Work. A Guide to Research and Writing. London: George Allen and Unwin.

Pickup, A. J. and Anthony, W. S. (1968). 'Teachers' marks and pupils' expectations: the short-term effects of discrepancies upon classroom performance in secondary schools'. British Journal of Educational Psychology, 38 (3) pp.302-309.

Pilliner, A. E. G. (1968). 'Examinations': In Butcher, H. J. (Ed.): Educational Research in Britain: London University Press. pp. 167-184.

Popham, J. W. (1975). Educational Evaluation: Prentice Hall, Inc; Englewood Cliffs, N.Jersey.

Popham, J. W. (1981). Modern Educational Measurement: Prentice Hall, Inc; Englewood Cliffs, N.Jersey.

Rowntree, D. (1977). Assessing students: How shall we know them? Harper and Row, Publishers.

Sagara, E. (1975). Girls and Boys: are there differences in ability? Paper presented at the Girls and Science Education Conference, Chelsea College.

Sands, M. K. and Bishop, P. E. (1984). Practical Biology. A guide to teacher assessment: Bell and Hyman. London.

Sands, M. K. and Forrest, G. M. (1981). 'A Scheme for teacher assessment of practical work in advanced level Biology and its moderation'. Journal of Biological Education, 15, (2) pp.146-150.

Schwab, J. J. (ed.) (1963). Biology Teachers' Handbook. BSCS Publication: John Wiley and Sons, New York.

Science Teachers Association of the United Republic of Tanzania. Annual Meeting 1972. Dodoma 1972 - (Mimeo).

Scriven, M. (1967). 'The Methodology of Evaluation' In Perspectives of curriculum Evaluation. Tyler et.al (eds.) AERA Monograph Series on Curriculum Evaluation No.I. Rand McNally, Chicago. pp.39-83.

Seka, B. (1980). An Evaluation of Secondary Mathematics. Institute of Education Dar-es-Salaam.

Shayer, M. (1974). 'Conceptual demands in the Nuffield 'O-level biology course'. School Science Review, 56 (195) pp.381-388.

Shayer, M. and Adey, P. (1981). Towards a Science of Science Teaching-Cognitive development and curriculum demand. Heinemann Educational Books.

Simpson, M. and Arnold, B. (1982). 'Availability of prerequisite concepts for learning biology at certificate level'. Journal of Biological Education, 16 (1) pp.65-72.

Smith, F. M. and Adams, S. (1972). Educational Measurement for the Classroom Teacher. Harper and Row, Publishers.

Stake, R. E. (1967a). Towards a Technology for the Evaluation of Educational Programs: In Perspectives of Curriculum Evaluation, Tyler et.al (Eds.). AERA Monograph Series on Curriculum Evaluation No.I. Rand McNally, Chicago, pp.1-12.

Stake, R. E. (1967b). 'The Countenance of Educational Evaluation'. Teachers College Record, 68, pp.523-540.

Stenhouse, L. (1975). An Introduction to Curriculum Research and Development, London, Heinemann.

Stufflebeam, D. L. (1967). 'Evaluation as Enlightenment for Decision making'. In Improving Educational Assessment, and An Inventory of Measures of Affective Behaviour'. Association for Supervision and Curriculum Development, Commission on Assessment of Educational Outcomes. Washington, D.C. pp.41-73.

Stufflebeam, D. L. et al. (1971). Educational Evaluation and Decision Making. Peacock, Itasca, Illinois, F. E. Peacock for Phi Delta Kappan National Study Committee on Evaluation.

Taba, H. (1962). Curriculum Development: Theory and Practice. Harcourt, Brace and World, Inc.

Tamir, P. and Glassman, F. (1971). 'A Practical Examination for BSCS Students: A progress report'. Journal of Research in Science Teaching, vol.8 (4) pp.307-315.

Tamir, P. and Glassman, F. (1972). 'The Practical Mode. A Distinct Mode of Performance in Biology'. Journal of Biological Education. (6) pp.175-182.

TANU, (1967). The Arusha Declaration. Government Printers. Dar-es-Salaam.

TANU (1974) The Musoma Resolution. Government Printers. Dar-es-Salaam.

TANZANIA, United Republic of, (1973). The Act of Parliament for the Establishment of the National Examinations Council of Tanzania (No.21 of 1973). Government Printers.Dar-es-Salaam.

TANZANIA, United Republic of, (1975). The Act of Parliament for the Establishment of the Institute of Education (No.13 of 1975). Government Printers. Dar-es-Salaam.

Thorndike, R. L. and Hagen, E. (1969). Measurement and Evaluation in Psychology and Education: John Wiley and Sons. Inc. (3rd Ed.).

Thyne, J. M. (1974). 'Types of questions'. In Principles of Examining: University of London Press. pp.173-184.

Tuckman, B. W. (1972). Conducting Educational Research. Harcourt Brace Jovanovich, Inc.

Turney, B. and Robb, G. (1971). Research in Education: An Introduction. The Dryden Press Inc. Hinsdale, Illinois.

Tyler, R. W. (1931). 'Techniques for Evaluating Behaviour'. Educational Research Bulletin, Vol.13. pp.1-11.

Tyler, R. W. (1931). 'A Generalized Technique for Conducting Achievement Tests'. Educational Research Bulletin, Vol.10. pp.199-208.

Tyler, R. W. (1949). Basic Principles of Curriculum Instruction. University of Chicago Press.

Tyler, R. W. (1967). 'Changing Concepts of Educational Evaluation'. In Perspectives of Curriculum Evaluation: Tyler et.al (Ed.). AERA Monograph Series on Curriculum Evaluation No.I. Rand McNally, Chicago. pp.13-18.

Tyler, R. W; Gagné, R. and Scriven, M. (Eds.). (1967). Perspectives of Curriculum Evaluation: AERA Monograph Series on Curriculum Evaluation, No.I. Rand McNally, Chicago.

UNESCO (1966). Improving the Teaching of Science. A draft paper for the 5th Session of the Economic and Social Council's Advisory Committee on Science and Technology to Development. UNESCO: Paris.

Ward, C. (1980). Designing a scheme of Assessment: Stanley Thornes (Publishers) Ltd.

Welford, G; Harlen, W. and Schofield, B. (1985). Practical Testing at Ages 11, 13 & 15 Department of Education and Science, Welsh Office.

Wheeler, D. K. (1967). Curriculum Process. University of London Press.

Whittaker, R. J. (1974). 'The Assessment of Practical Work'. In Macintosh (Ed.). Techniques and Problems of Assessment. Edward Arnold (Publishers) Ltd. London.

Willmott, A. S. and Hall, C. G. W. (1975). O-level Examined: The Effects of Question Choice: Schools Council Research Studies. Macmillan Education. London.

Willmott, A. S. and Nuttall, D. L. (1975). The Reliability of Examinations at 16+ Schools Council Research Studies. Macmillan Education.London.

Wiseman, S. (Ed.). (1961). Examinations and English Education. Manchester University Press. pp.1-26.

Young, J. A. (1970). 'Continuous Assessment - a mixed blessing'. Education in Chemistry, 7 pp.9-13.

Youngman, M. B. (1979). Analysing Social and Educational Research Data. McGraw-Hill Company (U.K.) Ltd.

APPENDIX I

UNITED REPUBLIC OF TANZANIA

MINISTRY OF NATIONAL EDUCATION

VOL. 2 - SECONDARY SCHOOL

SCIENCE SYLLABUSES - 1976

THE BIOLOGY SYLLABUS

FORM I - IV

BIOLOGY SYLLABUS FORM 1-1V

Introduction:

Knowledge of Biology, it must be emphasized, is a cultural requirement, for we need to know ourselves and our environment. Since all man's needs find their satisfaction in nature, it is through a better knowledge thereof, that he will be able to raise his standard of living. Application of Biological principles concepts and knowledge should be the spur in Biology courses, since knowledge for which no use can be perceived is soon forgotten, and cannot to a worthwhile education.

Teaching of Biology tries to look particularly into its role in agriculture: animal husbandry, crop husbandry, forestry, fisheries, nature conservation, para-agricultural industries, human nutrition, heredity, ecology, and the interaction of man and his environment: so that man could utilize his natural heritage (land, plants and animals) to his best advantage without destroying them.

The main objective in teaching biology, therefore, is to enable the graduates in biology to apply the biological principles, concepts, knowledge and skills so acquired, to their rural community, so that they could best utilize their natural heritage (land, plants and animals) without destroying them, to raise their standard of life and free themselves from diseases, poverty and ignorance.

TOPICS:	NOTES:
1. INTRODUCTION TO BIOLOGY: - Meaning - Objectives - Methods.	
2. INTRODUCTION TO PRINCIPLES OF CLASSIFICATION: Diversity of living things and their distribution.	Characteristics of living things, living and non-living things compared. Plants and animals compared. Sorting of living things based on similarities and differences, habitats, sizes and shapes by inquiry approach.
Major groups of Animals and Plants. One celled organisms Protozoa: Trypanosome/ Plasmodium/Amoeba/Paramecium.	Observation of External features, habitat and movement.
Worms: Ascaris, earthworm and Tapeworm. Arthropods: grasshopper, mosquito, crabs,	Use of pertinent features to group animals into vertebrates and invertebrates; plants into flowering and non flowering.
Vertebrates: fish, amphibia, reptilia, birds, mammals.	Emphasize general features without going into details of anatomical and physiological aspects. Practical observation of different types of feathers.

TOPICS	NOTES:
3. CELLS STRUCTURE AND FUNCTION:	.
Structure and function of Plant and animal cells.	Structure of typical Plant and animal cells. Cell structure to include only cell wall; Plasma membrance, cytoplasm, nucleus and cell vacuole. Comparison between Plant and animal cells.
Plant and Animal tissues, organs and systems: e.g. Vascular, Ground, Mechanical & Connective.	Stress how structure (shapes of various tissues, organs and system) relates to function.
Interaction of materials within and around cells.	Diffusion & Osmosis, turgidity, plasmolysis, wilting. Permeability in the living cell. (Practical experiments to illustrate the above phenomena using living and non living materials) - stress the differences.
4. PARASITES AND DISEASES OF MAN:	
Meaning and survival problems of parasites.	
Infections of Biological origin. Principles concerning the relationship between man and the pathogenic agents.	
The world of microbes: Experimental introduction to techniques and methods of micro-biology.	Microscopic observation of bacteria. Culturing of bacteria.
Diversity and characteristics of pathogenic agents.	Refer to Bacteria, viruses and worms and insects as vectors, life cycle of mosquitoes.

TOPICS:	NOTES:
Major endemic diseases in Tanzania.	Refer to Bilharzia, typhoid, Malaria, leprosy, T.B., dysentery and Venereal diseases.
Principal infectious diseases and their classification according to the methods of transmission.	Infection transmitted by air, water, food, contagion from the ground, contact & vectors. Control of such diseases.
5. SOIL AND PLANT GROWTH:	
Soil formation - agents of soil formation Chief constituents and properties of soils. viz. living components and non living components.	Soil forming agents to include:- Running water, wind, temp, dilute acids, animals and plant roots. How soil properties are related to soil forming agents. Measurements of the following mineral contents of soil N.P.K. and lime using soil test kit.
Soil Profile	To illustrate natural layering of soils, during field work.
Soil analysis by Mechanical process (particle size)	Use of sieve and sedimentation
Soil water; Measurement of soil water Content:- Water retentivity, capillarity, porosity, and field capacity.	Origin of water:- rainfall, irrigation, atmospheric humidity: Compare soil water contents in sandy, loam and clay soils. (practicals necessary)
Soil fertility and plant growth. Loss of soil fertility and control measures. Manures and fertilizers:types and their uses.	Factors leading to loss of soil fertility to include - leaching, plants, erosion, wind, poor farming practices, burning. Control measures to include; control of soil erosion, reafforestation, application of fertilizer, guard against excessive over grazing & good farming practices. Demonstration plots are very essential.

TOPICS:	NOTES:
6. NUTRITION:	
(a) Plant Nutrition Leaf structure; gaseous exchange.	Microscopic observation of T.S. of a named leaf to identify the different Parts/layers: Distribution of stomata, chloroplasts and air spaces. A variety of leaves could be observed.
Water and mineral absorption in plants.	Water uptake by roots; water culture experiments:- emphasize N.P.K.
Photosynthesis and conditions affecting it.	Raw materials and their sources. Simple treatment of the role of each constituent parts. Simple treatment of the mechanism of photosynthesis. And products of photosynthesis and their uses.
Synthesis of fats and proteins from carbohydrates.	
Food storage and storage organs.	
Special modes of nutrition in plants.	Saprophytism in mucor/rhizopus: parasitic plants:- striga/ orabanche on tabacco/mistletoe. Adaptations of these modes of nutrition as a result of lack of chlorophyll or roots.
(b) Animal Nutrition. Food sources and habits for human and other animals.	Omnivorous, carnivorous and herbivorous distinguished.
Basic food materials - viz. Carbohydrates, fats, proteins; their simple identification.	Include the sources of these food types. Reducing and non-reducing sugars.

TOPICS:	NOTES:
The role of constituents of food including water, mineral salts and vitamins.	Functions of each to the body including deficiency symptoms.
Balanced diet for animals.	Meaning; survey of food values and nutritional levels. Energy requirements for human beings. Food diets for expectant mothers (pregnant women) and babies during suckling and after weaning.
Malnutrition in Tanzania.	Meaning of malnutrition, deficiency disorders and their symptoms. Spread and causes of malnutrition in Tanzania.
Feeding and digestion in mammals.	
Adaptation of mouth parts in relation to types and methods of food uptake.	Tooth structure, (canines, molars, incisors); beaks. Dental formulae, teeth care and tooth decay.
The alimentary canals of man and cow in relation to food types.	Ruminants and monogastric compared.
Digestion process in man.	Functions of digestive enzymes of the mouth, stomach, duodenum, ileum. Functions of the liver.
Absorption and assimilation.	
Food preservation and storage.	Food contamination, the need and Methods of preservation.
ENZYMES and their characteristics	Simple experiments to demonstrate the effects of temperature, pH. enzyme concentration on enzyme activity.
7. TRANSPORT IN PLANTS AND ANIMALS:	
Structure of stems and roots.	Refer to monocots and dicots.

TOPICS:	NOTES:
Transpiration and conditions affecting its rate.	Experimental demonstration of conditions affecting transpiration, transpiration pull, root pressure.
Transportation of water in plants.	Use of coloured water to show the path of water from the roots to the leaves.
Composition and function of mammalian blood. Blood clotting.	Observation of mammalian blood under the microscope. Anaemia, Leukemia, Haemophilia.
Structure of heart and blood vessels. Blood circulation.	Observation of major blood vessels and lymph vessels. Causes of blood pressure.
8. BREATHING AND RESPIRATION:	
Organs of gaseous exchange. Breathing mechanism in man, insect and fish.	Reference to be made to the lung, trachea, gills and skin. Hazards of smoking.
Respiration - Breaking down of food to release energy. Glycolysis and role of A T P.	Need for energy - A T P as a store of energy, formation of A T P., Breakdown of A T P to A D P., muscle contraction and A T P break down. Accumulation of lactic acid.
Aerobic and Anaerobic respiration.	Formation of alcohol. Yeast in relation to bread & alcohol formation.
9. REGULATION IN ANIMALS AND PLANTS:	
Excretory Organs.	Formation of urea in mammalian liver.
Formation of excretory products of animals.	
Elimination of excretory products. Factors affecting amount of excretory products.	Functions of kidney - uriniferous tabules osmoregulation - Skin as excretory organ.
Temperature regulation in warm and cold-blooded animals.	The need for temperature regulation; hibernation and torpidity in cold blooded animals.

TOPICS:	NOTES:
Excretion in Plants.	
10. CONTROL AND CO-ORDINATION SYSTEMS IN ORGANISMS:	
Response in plants and animals to stimuli-light, water and gravity.	Hormonal explanations of these responses.
Plant hormones (especially auxin)	Application of plant hormones as weed killers.
Nervous System	Structure of the nerve cells (Neurons) Types of nerve cells i.e. motor, sensory and relay. Nerves: Sensory, motor and mixed.
Spinal cord.	Spinal cord as seen in T.S. Simple and conditioned reflex.
Brain.	Gross structure (viz. olfactory lobe, cerebelum, cerebrum medulla oblongata) Functions of these parts. Demonstrations of brain parts to be carried out using models.
Sense Organs	Parts of the eye and their functions. Formation of image.
eye	Eye defects and their corrections.
ear	Detection of sound and balancing mechanism.
skin	As a sense organ of touch, pressure and temperature.
tongue	Location of taste buds (viz. sweet, sour, bitter and salt).
nose	As an organ of smell (olfactory).
Endocrine system: Adrenals, pancreas , pituitary, thyroid, ovary and testis. Stomach duodenum.	Location of position, secretions and functions. Emphasize growth hormones. Thyrotropic hormones. Gonadotropic hormones, thyroxin, oestrogen,

TOPICS:	NOTES:
	testosterone, adrenalin, secretin and gastrin.
11. MOVEMENT IN ANIMALS:	
Simple movement in Amoeba.	
Muscular movement.	How muscle and bones bring about movement.
Skeleton and joints of limbs.	Antagonistic muscular activity. Details of specific bones not required. Mechanism of limb movement.
Movement in insects, fishes, birds and mammals.	Structural adaptations of each to movement in water (fish), air (birds) and land (mammals).
12. GROWTH:	
Meaning of growth.	Change in size and form.
Factors affecting growth.	Emphasize Food (nutrition), environment, hormones.
Growth in plants: seed structure and germination: essential conditions for germination.	Part played by each seed part, types of germination (Epigeal hypogeal). Experimental determination of conditions necessary for germination.
Regions of growth in plant roots and stems.	Experiments to show the growing regions of shoot and root.
Mitosis.	Duplication of chromosomes, details of specific stages not expected. Microscopic observations of root tip cells.
Primary and Secondary growth in plants.	Formation of rings, comparison between dicot. and monocot. Stems and roots.
Growth in animals: Metamorphosis in Frog:	In frog metamorphosis exclude embryological details - distinguish

TOPICS:	NOTES:
house fly and cockroach or grasshopper.	complete and in-complete metamorphosis; include life cycles.
13. REPRODUCTION:	
Meaning and importance of reproduction.	Asexual reproduction in Mucor/Rhizopus, yeast.
Asexual reproduction in Plants.	Bryophyllum - vegetative propagation.
Sexual reproduction in flowering plants. Asexual and sexual reproduction compared.	Practical study of named flowers to illustrate characteristic of insect and wind pollinated flowers. Cross and self-pollination.
Seed and fruit formation.	Types of seeds and fruits.
Dispersal of seeds.	Practical observation of types and methods of seed dispersal.
Sexual reproduction in higher animals.	The following animals to be studied - Fish (Tilapia). Birds (weaver bird or chicken) and a small mammal study to include breeding habits, incubation period and parental care. Experimental observations on reproductive organ to be carried out.
Types of fertilization including mating.	Reproduction in Man should be discussed.
Nutrition of embryo and the young.	
14. GENETICS:	
Cell divisions - Mitosis and Meoisis.	Mitosis and Meiosis in relation to DNA replication: DNA molecule simply treated.
Chromosomes and concept of genes.	
Mendels laws of inheritance.	Concept of Dominance and recessive (refer to monohybrid crosses only).

TOPICS:	NOTES:
Mutations and gene action.	Meaning of mutation.
Methods of breeding in plant and animals of a harder and better characteristics.	Selective breeding in maize, cattle and poultry to be discussed.
Variation among plants and animals of the same or different species.	Include sources of variation - genetical and environmental simply treated.
Blood groups and tran- sfusion.	Refer to A.B.O. blood types and Rh-factor.
15. EVOLUTION:	
Theory of evolution. Over- production; struggle for existence; individual variation, survival of the fittest.	Adaptation and mutations as agents of evolution.
16. INTERDEPENDENCE OF ORGANISMS:	
The concept of ecosystem.	Plants as producers; animals as consumers; different types of consumers.
Food chains and food webs.	
Balance of Nature.	Nitrogen cycle, carbon cycle.
Insects as pests of crops.	Life cycle of maize stalk borer/ cotton stainer and weevil/coffee mealybug and their effects to crops.
Parasites of animals.	Refer to ticks, fleas and lice.
Organisms and their succession.	A study of plant succession or animal succession by transect method.
Population and communities: Population growth and control.	

TOPICS:	NOTES:
Factors affecting distribution of plants and animals.	Refer even and uneven distribution.
17. MAN AND HIS NATURAL RESOURCES:	
The importance of natural resources.	Land, plants and animals.
Conservation of the natural resources.	Wildlife, Forests and Water.
Pollution of the environment - water, air and household.	Effects of pollution on human life.

APPENDIX II

TEACHER'S GUIDE

ON

'O' LEVEL BIOLOGY SYLLABUS

BIOLOGY PANEL

INSTITUTE OF EDUCATION

SEPTEMBER 1978

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
1 INTRODUCTION TO BIOLOGY a. Meaning b. Objective c. Methods	1. Enable pupils to identify biology as a science that deals with the environment. 2. To enable pupils to learn scientific methods such as observation, questioning, hypothesizing, experimenting and drawing conclusions.		
2 INTRODUCTION TO PRINCIPLES OF CLASSIFICATION A Diversity and distribution of living organisms. 1 Characteristics of living organisms. 2 Comparison of living and non-living things.	1. Pupils should be able to distinguish plants from animals using morphological features. 2. Using morphological features, pupils should be able to place animals and plants into major groups.	1. Pupils to collect a variety of living organisms (each teacher to make sure they have a good cross section of animals and plants). 2. Pupils with the guidance of a teacher to sort out the organisms collected and place them into their respective groups using morphological features.	A modern Introduction to Biology by Beckett. Introduction to biology a new tropical edition by Mackean.

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
3 Introduction to the identification and classification of major plant and animal groups on the basis of similarities and differences, habitats, sizes and shapes through inquiry approach.	3. Pupils should be able to enumerate factors controlling the distribution of living organisms.	(Presence of hair, feathers, leaves, flowers etc.). Pupils to record the features of the habitats from which organisms are collected.	
	4. Pupils should be able to enumerate and describe all life processes characteristic to living organisms.	3. Pupils to observe various activities performed by living organism e.g. a. Breathing (observe the taking in of oxygen and release of carbondioxide). b. Growth (Measure the height of a seedling every day for a week, weigh a young, small animal every day for a week.) c. Excretion. Observe elimination of excretory materials e.g. sweating, urination.	New biology for tropical Schools by Stone and Cozens.

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
		d. Movement (observe movement in insects and other small animals). e. Irritability (Pupils to observe a response of an arm when pricked by another pupil) f. feeding (Observe a small animal when supplied with food)	Modern biology for Secondary Schools by Pereira and Ramalingan. - Beckett - Mackean - Stones & Cozens
	Pupils should be able to distinguish living organisms from non-living things on the basis of life processes.	Pupils to observe and discuss the following life processes: Nutrition,movement, reproduction and growth.	-do-
B. Major groups of plants 1. Flowering Plants a. Seeds produced and covered in the ovary (Angiosperms) i. Characteristics of seeds. ii. Comparison of dicots and monocots.	Pupils should be able to: 1. distinguish flowering plants from non flowering plants. 2. Draw and label different parts of a flower. 3. Distinguish monocots from dicots.	1. Pupils to collect and observe the parts of a dicotyledon 2. Pupils to compare the seed structures of dicot with those of a monocot plant e.g. Maize, rice, wheat for monocots and beans, castor oil and groundnuts for dicots.	-do-

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
b. Seeds produced but not covered in the ovary (Gymnosperms) only a mention.		3. Pupils to compare leaves and rooting systems of dicots and monocots.	
2. Non-flowering plants a. Algae b. Fungi c. Mosses d. Ferns Only Introductory account on each group is needed Life cycle studies are not required	1. Pupils should be exposed to different types of non flowering plants. 2. Pupils should be able to describe the methods by which non flowering plants reproduce.	Pupils to collect, draw, label and group the following specimens: moulds (from bread, rotting fruits, cow dung etc.) Algae (e.g. Spirogyra, ulva, Ficus) Ferns: Mosses.	- do -
C. Major animal groups. 1. Invertebrates (Animals without backbone) a. Protozoa e.g. Amoeba Paramecium, Plasmodium and Trypanosoma.	1. Pupils should be able to: group animals into their phyla on the basis of external features. 2. Compare invertebrates and vertebrates 3. Describe distinctive characteristics of each groups.	1. The teacher to demonstrate to the pupil how a microscope is used. 2. Pupils to observe prepared slides of protozoa. 3. Pupils to observe live specimen under the microscope.	

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
b. Metazoa i. worms e.g. roundworm, earthworm and tapeworm. ii. Arthropods. e.g. Insects-grasshoppers, mosquitoes, flies Crustacea -crabs, prawns Arachnida-scorpions, spiders, mites Chilopoda - centipedes Diplopoda- millipedes	4. Use a microscope to observe small animals such as protozoa. 5. Separate classes of Arthropoda using number of legs.	(If it is difficult to obtain some of the specimens contact nearest hospital for <u>Plasmodium</u> and <u>Trypanosoma</u>. 4. Obtain live and/or preserved specimens of metazoa and study their morphological features. NB. See advancement of these from protozoa. 5. Observe external features of arthropods especially appendages.	- do -
2. Vertebrates (animals with backbone) e.g. Fish, amphibians, reptiles, birds & mammals.	1. Pupils should be able to classify vertebrates into the five classes.	1. Observe and study the external features of each class. 2. list down the distinctive features of each group (class).	- do -

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
3. CELL STRUCTURE AND FUNCTION A structure of cells as seen under light microscope. 1. nucleus 2. Cell wall 3. Cell membrane (especially animal cells) 4. Cytoplasm 5. Cell vacuole	1. Pupils should be able to draw and label various parts of a cell.	1. Pupils to study the parts of a microscope and their functions. 2. Pupils to observe, draw various plant and animal cells.	- do -
B. Comparison between plant and animal cell.	Pupils should be able to distinguish animal cells from plant cells.	to observe and draw various plant and animal cells.	- do -
C. Functions of 1. Nucleus 2. Cell wall 3. Cell membrane 4. Cytoplasm 5. Cell vacuoles These can be taught in relation to A after the pupils have under-stood the structure of a cell.	Pupils should be able to: 1.describe the functions of the nucleus, cellwall, cell membrane, cytoplasm and cell vacuoles. 2.Relate the function with the structure.	- ditto -	- do -

TOPIC OUTLINE	OBJECTIVE	SUGGESTED PUPIL ACTIVITY	REFERENCES
D. Organization of Cells into tissues. Under this topic consider: 1. Muscular tissue 2. Connective tissue e.g. blood. 3. Vascular tissue in plants.	Pupils should be able to: 1. distinguish the structure of one tissue from another. 2. Relate the function of a tissue to its structure.	Observe and draw the following tissues under the microscope. a. muscle b. blood c. vascular bundles.	- do -
E. The organization of tissues into organs; for example. 1. The mammalian heart (has muscular tissues and connective tissues) 2. A leaf of a plant (has epidermal, palisade and vascular tissues) Emphasize function to cellular organisation.	Pupils should understand that organs are made up of tissues.	To examine charts, models or prepared sections of various organs.	- do -
F. Cell-water relation explain and demonstrate 1. Osmosis 2. Turgidity 3. Plasmolysis 4. Wilting 5. Diffusion	Pupils should be able to describe the process of water movement into and out of a cell and the effects that follow.	Perform experiments to demonstrate these phenomena.	- do -

TOPIC OUTLINE	OBJECTIVE	SUGGESTED PUPIL ACTIVITY	REFERENCES
4. SOIL IN RELATION TO PLANT GROWTH. A. Soil formation. 1. Concept of soil formation. 2. Soil forming agents to include running water, wind, temperature, dilute acids, animals and plant roots.	1. Pupils should be able to explain how soil is formed from the parent rock by various methods.	1. Field excursions to observe and record soil forming agents in action.	- do -
B. Chief constituents and properties of soils. 1. Living and non living components of soil. 2. Mineral composition of soil. 3. How soil properties are related to the nature of soil formation	Pupils should be able to: 1.enumerate the various components of soil. 2.Discuss the properties of soil in relation to their constituents. 3.Relate soil properties to soil forming agents.	1. Soil analysis by mechanical process i.e. use of sieve and sedimentation. 2. Use of soil kit to determine mineral content of soil. 3. Practicals to determine constituent of soil e.g. air, mineral matter and living organisms.	- do - - do -
C. Soil profile - meaning of soil profile and how it is formed.	Pupils should be able to discuss how different layers of a soil profile are formed.	Observe a vertical section through soil and record natural layering of soils.	- do -

TOPIC OUTLINE	OBJECTIVE	SUGGESTED PUPIL ACTIVITY	REFERENCES
D. Soil and water 1. Origin of water in the soil (rainfall, irrigation, atmospheric humidity. 2. Water retentivity.	Pupils should be able to: 1. enumerate various sources of water in the soil. 2. Know the effectiveness of various types of soils in retaining water.	Measurement of water retentivity, capillarity, porosity and field capacity in various types of soil (e.g. loam, sand and clay)	-do -
E. Soil fertility and plant growth. 1. Meaning of fertility 2. Factors leading to loss of soil fertility. (Leaching, plants, erosion wind, poor farming practices, overgrazing and burning) 3. Control measures a. control of soil erosion. b. Reafforestation c. application of manures and fertilizers (say what it does).	Pupils should be able to: 1. describe how soil looses its fertility. 2. describe different ways applied in maintaining or restoring soil fertility.	1. field excursions to observe and record factors leading to loss of soil fertility. 2. Demonstration of plots under good and poor farming practices (pupils to be directly involved in the preparation of these plots)	- do -

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
d. guard against excessive over-grazing. e. good farming practices.			
5. NUTRITION A. Plant nutrition 1. leaf structure: examine morphology and anatomy of a leaf.	Pupil should be able to locate and discuss the distribution of stomata and chloroplast within and between leaves.	1. Microscopic examination and drawing of upper and lower epidermis of a dicot and monocot leaf to compare the distribution of stomata. 2. Microscopic examination and drawing of T.S. of a named leaf	- do -
2. Water and mineral absorption. a. water uptake by roots. b. Water culture experiments (N.P.K.)	Pupils should be able to: 1. discuss movement of water and minerals. 2. detect the deficiency symptoms (i.e. N.P.K.)	1. Carry out an experiment to demonstrate water uptake by roots using a coloured liquid. 2. Carry out a water culture experiment to find out the importance of N.P.K. to the plant.	- do -
3. Photosynthesis a. Gaseous exchange b. Light reaction: role of chlorophyll c. Dark reaction: fixation of carbon:	Pupils should: 1. know that green plants manufacture their own food. 2. describe the process of photosynthesis	Carry out experiments to show: 1. that Co_2 is necessary during photosynthesis. 2. that chlorophyll is necessary for photosynthesis.	

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
end products of photosynthesis.		3. Starch test.	
4. Synthesis of fats and proteins from carbo- hydrates, composition of carbohydra tes,fats and proteins.	Pupils to be aware of the fact that fats and proteins are synthesised in the plant through carbohydrates conversion.		- do -
5. Food storage and various storage organs.	Pupils should be able to locate and identify food types stored in different plant parts.	To collect different plant storage organs, observe and draw them.	
6. Special modes of nutrition in plants. a. Symbiotic relation- ships b. Saprophytism (Mucor/ Rhizopus) c. Parasitism (e.g.Mistle- toe) Stress adaptations of plants to these modes of nutrition.	Pupils should be able to: 1.Understand that besides photosythesis there are other methods of nutrition. 2.describe adaptation to these modes of nutrition.	1. Collection of leguminous root nodules. 2. Culturing mucor on bread. 3. Locate and observe mistletoe.	- do -
B. Animal nutrition. 1. Food sources for a. herbivores b. canivores c. omnivores	Pupils should be able to 1. distinguish the three types of animal nutrition. 2. relate the type of nutrition to dentition.	1. Observe the teeth of herbivores, omnivores and record the adaptive features. 2. Draw the structure of a tooth.	- do -

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
2. Adaptations to these modes of nutrition. a. Types of teeth (homodont and heterodont dentition) b. Structure of a mammalian tooth. c. Types of beaks in birds.	3. relate the structure of different parts of the tooth to their functions.	3. Observe different types of beaks.	
3. Basic classes of food materials Discuss carbohydrates, fats and proteins, water, minerals and vitamin. a. Their identification in food stuffs (i.e.carbohydrates, fats and proteins) b. Their role in the human body.	Pupils should be able to: 1. enumerate the various components of food. 2. describe the various functions of foods stuffs in the human body. 3. Identify carbohydrates, fats and proteins in the food material.	Pupils to collect various food materials used locally and identify the various food classes they contain.	- do -
4. Balanced diet for man a. Meaning of balanced diet b. Survey of food values c. Malnutrition in Tanzania.	Pupils should be able to: 1.describe what constitutes a balanced diet.	1. Collect the various food stuffs used locally and determine their food content.	

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
i. meaning and causes ii. Deficiency disorders and their symptoms d. food diets for expectant and nursing mothers. e. foods for babies, during suckling and after weaning.	2. discuss the effects of not eating a balanced diet (deficiency symptoms and disorders). 3. choose the right food stuffs for expectant and nursing mothers. 4. advise their parents the right food for babies during suckling and after weaning.	2. Collect the various food stuffs used locally or used at school and discuss whether they constitute a balanced diet. 3. Observe deficiency symptoms in a local dispensary.	- do -
5. Digestion in animals a. Alimentary canal of man (non-ruminant) and a cow (ruminant) b. Digestive process i. Under this topic discuss the breakdown of food in the mouth,	Pupils should be able to: 1. distinguish ruminant from non ruminant digestive system. 2. Relate the digestive system to the type of food and feeding habits. Pupils should be able to: 1. explain the digestion process from the mouth up to the intestine.	To collect and observe models or preserved specimens of different digestive systems. Pupils to carry out experiments to illustrate:-the digestion of starch into simple sugars.	- do -

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
stomach, duodenum and intestine. ii. Enzymes and their characteristics. iii. Effect of temperature p^H, and enzyme concentration on enzyme activity. Use named enzymes to illustrate these effects. iv. functions of the organs related to digestive systems. e.g. liver, Pancreas, salivary glands.	2. discuss factors affecting enzyme activity. 3. relate the factors affecting enzyme activity to the digestion processes taking place in the mouth, stomach duodenum and intestine. 4. appreciate the importance of the liver, P and salivary gland in digestion.	2. Digestion of complex sugars into simple sugars. 3. The effect of p^H on Enzymes activity. 4. The effect of temperature on enzyme activity. 5. the effect of enzyme concentration on enzyme activity. 6. The effect of substrate concentration on enzymes activity.	
6. Absorption and assimilation	Pupils should be able to explain how end products of digestion are absorbed and later take part in metabolism of the body.		- do -

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
7. Food preservation and storage a. Storage of enzymatic activity. b. Need and methods of preservation using local methods: i. drying ii. Smoking iii. Salting	Pupils should be able to: 1. describe the various methods of food preservation and storage. 2. Discuss the merits and demerits of various methods of food preservation and storage. 3. Apply the appropriate method of food preservation and storage to different food materials.	1. Observe local methods of food preservation. 2. Observe industrial methods of food preservation e.g. canning, bottling, deep freezing etc.	- do -
6. TRANSPORT IN PLANTS AND ANIMALS a. Internal structure of roots and stems of monocots e.g. maize and dicots e.g. Broad Bean (<u>Vicia faba</u>) Emphasize vascular structures.	Pupils should be able to: 1. differentiate the vascular system of dicots from that of the monocots. 2. explain how water gets into the plant from the soil.	1. To examine and draw T.S. of monocot roots and stems and dicot roots and stems. 2. To carry out an experiment to demonstrate transpiration, root pressure, transpiration, pull and factors affecting transpiration.	

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
b. Osmosis The process of Osmosis: i) Movement of solvent and solute molecules across a cell membrane. ii) Movement of water from the soil into the root hairs, cortex, and ultimately into the vascular system. c. Transpiration and conditions affecting its rate. i) Transpiration pull. ii) Root pressure. d. Transport of water in plants.	3. describe the movement of water from the soil to the leaves and the forces involved.	3. To carry out an experiment using coloured water to show the path of water from the toots to the leaves.	
e. Composition and function of mammalian blood. i. Blood cells: red cells white cells and platelates ii. plasma, nutrients, hormones etc. iii. Blood clotting: process and its role.	Pupils should be able to: 1. describe the various components of blood and explain the function of each constintuent parts. 2. explain the process of blood clotting and its importance.	1. To observe under the micro-scope the various blood cells. 2. To dissert and draw the heart of a mammal. 3. Use models to study blood circulation.	- do -

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
f. Structure of heart and blood vessels i.e. heart chambers, arteries, veins and capillaries.	3. draw and label the main blood vessels of the heart & the four chambers of a mammalian heart.		
	4. distinguish arteries from veins structurally.	3. To observe under the microscope how blood flows using a tail of a tadpole or the web of mature frog or toad.	
g. Blood circulation i)Major arteries distributing blood to the rest of the body (from the heart). ii)Major veins returning blood to the heart.	5. Draw and label the main arteries from the heart and show the direction of blood flow.		
h. Lymphatic system i.Lymph ii.Lymphatic vessels and nodes. iii.function of lymphatic system.	6. distinguish composition and function of lymphatic system from blood circulatory system.	4. Locate lymph nodes in the human body.	- do -
7. BREATHING AND RESPIRATION A: Gaseous exchange in animals 1. Meaning of breathing (gas exchange).			

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
2. Morphology Anatomy, and function of organs of gaseous exchange. a. Epidermis b. Tracheal system c. gills d. lungs e. skin and tongue 3. Mechanisms of gaseous exchange.	Pupils should be able to: 1. relate the structure of organs of gaseous exchange to their function. 2. explain how the organs of gaseous exchange have been able to increase the surface area at the same time minimized volume.	Pupils to observe mechanisms of gaseous exchange in insects, fish, frog/toad and man. 2. Pupils to draw organs of gaseous exchange. (i.e. tracheal system, lungs and gills).	- do -
C. Gaseous exchange in plants. Gaseous exchange through 1. stomata 2. lenticels 3. roots D. Hazards of imparing gaseous exchange 1. smoking 2. water-logging etc.	Pupils should be able to: 1. explain how gaseous exchange takes place through stomata, lenticels and roots. 2. relate the structure of these organs to their function. 1. Pupils should be aware of the hazards related to smoking and water-logging.	1. Pupil to observe and draw the organs of gaseous exchange in plants (i.e. stomata lenticels and roots) Pupils to demonstrate how smoking and water logging interfere with gaseous exchange.	- do - - do -

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
E. TISSUE respiration Oxidation of glucose and fatty acids in the cell to release energy (ATP); carbon dioxide and water are also released as by products.	Pupils should be able to: 1.define respiration. 2.explain how energy is obtained from glucose.	Pupils should carry out an experiment to find out and detect the products and by products of respiration (e.g. heat-energy, water, carbon dioxide, alcohol).	
1. Aerobic respiration a. utilization of oxygen-Carbon dioxide water Energy(ATP) b. Utilization of ATP to provide energy in all biochemical process in all biochemical process in the body.	3.differentiate between aerobic and anaerobic respiration.		- do -
	4. explain the meaning of oxygen debt.		
2. Anaerobic respiration (fermentation) a.meaning (oxygen not utilized) b.accumulation of lactic acid during heavy exercise when oxygen	5.outline the application of fermentation in various industries.		

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
is inadequate (oxygen debt) c. fermentation Glucose Alcohol + CO_2 + Energy (ATP) d. Commercial application of fermentation process in brewing and baking.			
REGULATION IN PLANTS & ANIMALS A: Meaning of regulation B: Regulation in animals 1. Structure and functions of regulatory organs. a. Kidney: i. General structure of kidney in relation to urinogenital system ii. Nephrone as a structural and functional unit of the kidney. iii. Function of kidney as an osmoregulator iv. Factors affecting excretory products (temperature,water and hormones)	Pupils should be able to: 1. Define regulation. 2. draw and label the various parts of the mammalian kidney and its connection to the urinogenital system. 3. relate the functions of the kidney to its structure. 4. discuss the factors affecting the quantity and type of excretory products.	1. Pupils should observe the external features of the mammalian kidney and its position in the body. 2. Pupils should dissect the mammalian kidney and draw and label the internal structure. 3. Pupils should observe and collect the various excretory materials by fish, amphibians, reptiles, birds & mammals.	- do -

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
b. Skin i. Generalized structure of a mammalian skin.	Pupils should be able to: draw and label T.S. of a mammalian skin.	1. To observe, draw and label prepared T.S. of a mammalian skin.	
ii. Function: - elimination of water - elimination of salts and nitrogenous wastes - regulation of body temperature.	2. discuss the various functions of the mammalian skin. 3. relate the function of the skin to its structure.	2. Observe cooling effects of evaporating liquid on their skin.	- do -
C. Lungs i. General structure of the lungs. ii. Functions - elimination of carbondioxide and excessive water. D. Liver: External structure and its location in the body. ii.Functions. - deamination - detoxication - regulation of sugars.	Pupils should be able to: 1. draw and label the structure of the lungs. 2. discuss the function of the lung in relation to its structure. 3. point out the location of the liver in a mammalian body. 4. relate the function of the liver to its structure.	1. Exhale onto the glass to detect water. 2. Exhale through lime water.	- do -

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
E. Excretion in plants. Organs of excretion 1. Leaves a. Gaseous exchange through stomata. b. Transpiration (elimination of excess water). c. Leaf fall (elimination of metabolic wastes translocated to old leaves). 2. Bark. a. exudation of substance through the bark (tanins, resins, gums etc.) b. gaseous exchange through lenticels.	Pupils should be able to explain how plants remove excess gases, water and other products of metabolism.	1. Place a potted plant in a bell jar and observe moisture on the inner surface. 2. Locate trees with gums e.g. Acacia, pine, mangoe, etc.	- do -
9. CONTROL AND COORDINATION SYSTEM IN ORGANISMS A. Stimuli: perception and responses in organisms.	1. Pupils should appreciate that some changes in the environment bring about behavioural changes in organisms.	1. Drawing of the nerve cells, brain and spinal chord. 2. Pricking and blinking to demonstrate reflex action.	

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
B. Systems involved in bringing about control and coordination in animals. 1. - Nervous system a. structure of nerve cells and different types of nerves should be taught.	2. Pupils should appreciate that these behavioural changes are for survival purposes.	3. Observe and draw a Cow's/ goat's/sheep's eye and ear.	
b. Central nervous system. i. Brain Only the major morphological features and functions of cerebral cortex, medulla oblongata and cerebelum.	3. To be able to describe the structure of a nerve cell in relation to function. 4. To know how to draw and label different parts of a nerve cell.	4. Draw a tongue to show distribution of different taste buds.	- do -
ii. Spinal chord Morphological features and functions only.	5. To draw and label a cross section of the brain cut into two identical halves.		
2. Peripheral nervous system: a. Cranial nerves related to sense organs. b. Spinal nerves.	6. Pupils should be able to explain the function of the brain and the spinal chord.		
c. Sense organs. i. Eye - structure and function.	7. To explain how a reflex action is brought about.		

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
ii. Ear – structure and function iii. Skin as a sense organ of touch, pressure and temperature. iv. Tongue (taste buds) (structure and function)	8. To describe different sense organs in relation to their functions. 9. To be able to explain how each sense organ perceives stimulus.		
Hormonal system. a. Plants Consider auxins and their functions in relation to responses of plants to their environment i.e. Phototropism, hydrotropism, geotropism.	Pupils should be able to: 1. relate auxins to tropism. 2. Locate the position of these glands. 3. enumerate the hormones produced.	1. Carry out tropism experiments. 2. Dissect a mammal and locate the positions of the glands. 3. Observe the location of glands on charts or models.	- do -
b. Endocrine system in animals. Consider the following glands; adrenal pancreas, thyroid, pituitary, ovary and testes. Consider only 1. position, 2. the hormones they produce 3. functions of hormones mentioned above.	4. explain the function of these hormones. 5. explain disorder of over or under secretion.		

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
4. Show the results of their over or under activities. (Producing disorders e.g. gigantism, goitre, dwaftisms in case of thyroid gland-throxin).			
c. The integration of central nervous and the hormonal systems in control and coordination. Consider an animal in fright e.g. a man seeing a lion. Functional units-eye, brain, insulin, adrenalin, heart, muscular activity i.e. running or fright. eye - sees the animal brain - receives what is perceived and commands hormonal glands e.g. adrenal gland to produce adrenalin and so increase the heart beat and pancreas to produce insulin and so mobilize sugar for energy.	Pupils should be able to explain how the nervous system works in conjunction with the hormonal system.		- do -

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
muscular activities - running away. In the stomach and duodenum there are certain cells that produce hormones There are several examples to illustrate the intergration of nervous system with hormonal system, therefore you are not limited to the above example. 10 MOVEMENT A. Simple movement in protozoa. 1. ciliary movement (Paramecium) 2. flagellar movement (e.g. Euglena) 3. Pseudopodial movement e.g. Amoeba.	Pupils should be able to: 1. Describe how non muscular organisms such as protozoa are able to move. 2. distinguish the three types of simple movements from each other.	1. To observe movement in Amoeba, Paramecium and Euglena and represent these movements diagramatically.	- do -
B. Muscular movement 1. structure of muscle simply treated 2. Types of muscles (i.e. smooth muscle, striated muscles and cardiac muscles) simply treated.	Pupils should be able to: 1. describe the structure of muscles. 2. distinguish the three types of muscles from each other. 3. relate the structure to the function a muscle plays.	1. To observe prepared slides of the structure of smooth striated and cardiac muscles. 2. Pupils to collect different types of skeletal structures and preserve them. 3. Pupils to draw different types of joints.	- do -

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
C. Skeleton and joints of limbs. Endoskeleton and exokeleton. D. How muscles and bones bring about movement in animals.	4. discuss how muscles together with skeletons bring about movement in insects, fish birds and mammals.	4. Pupils to observe and draw muscle attachment to long bones (only muscles related to movement).	
Movement in: 1. fish 2. birds 3. mammals 4. frog/toad 5. snakes.	1. Pupils should be able to discuss structural adaptations of each organism to movement in water (fish) air (birds) and on land (Mammals).	1. Pupils to observe morphological features of fish, birds and mammals. 2. Pupils to observe and compare complete skeletons of fish birds and mammals.	- do -
11. GROWTH 1. Meaning of growth (Increase in protoplasm) 1. change in size, form and weight. 2. Factors affecting growth a. nutrition b. environment c. hormones.	Pupils should be able to: 1. define growth. 2. explain the role played by each factor in growth. 3. draw and label the different parts of a dicot and a monocot seed.	1. Pupils should carry out experiments using auxonometer or meter rule for measuring vertical growth. 2. To set long term experiments to verify effect of environmental factors. 3. Dissection of monocot and dicot seeds for observation and recording of parts.	- do -

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
3. Growth in plants a. seed structure b. part played by each seed part in germination. c. types of germination. d. conditions essential for germination.	4. discuss the role of each part of a seed in germination.	4. Experiments to determine conditions necessary for germination.	
4. Regions of growth in plants. The apical meristems of both stems and roots.	5. relate the positions of meristematic tissues to the type of growth.	5. To plant maize grains and bean seeds and observe the type of germination.	- do -
5. Mitosis a. meaning b. Importance (i.e.growth)	6. describe the process of mitosis and relate it to growth.	6. Experiment to show the growing regions of shoot and root.	
6. Primary and Secondary growth in plants. a. meaning b. Primary growth - increase in length. c. Secondary growth - formation of rings leading to increase in growth. Compare monocot and dicot stems and roots.	7. enumerate and explain the part played by each condition of germination. 8. tell the importance of secondary growth to the plant itself and to man.	7. Observation of mitotic division under the microscope from prepared slides. 8. Transverse section of both monocot and dicot stems and roots to show variations of secondary growth. 9. To measure primary growth and secondary growth.	

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
	9. explain the economic importance of secondary growth.	10. Estimate the age of a tree using anual rings.	
7. Growth in animals. Growth curves of common animals. e.g. fish, amphibian (frog) insects, birds, reptiles and mammals.	Pupils should be able to: 1. interprate growth curves.	Compare growth curves of different animals as pupil project.	
a. Distinguish between complete and incomplete matamorphosis in insects.	2. distinguish between complete and incomplete metamorphosis in insects.	To study the life cycles of insects and note the differences between complete and incomplete metamorphosis.	- do -
b. Life cycles of housefly and cockroach or grasshopper.	3. describe the life cycles of a housefly and a cockroach or grasshopper.	To study metamorphic stages in a frog/toad.	
c. Metamorphosis in frogs excluding embryological details.	4. outline the major metamorphic stages of frogs.		
12. REPRODUCTION	Pupils should be able to:		
1. Meaning and importance of reproduction.	1. discuss the importance of reproduction.		
2. Asexual reproduction in plants a. Natural vegetative propagation i. budding e.g. yeast	1. name and describe the different types of reproduction.	1. Examine organisms which show different types of reproduction.	- do -

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
ii. sporulation e.g. Mucor iii. binary fission e.g. Amoeba iv. suckers e.g. banana v. leaves e.g. Bryophyllum vi. stolons vii. rhizomes viii. bulbs, e.g. Onions ix. corms x. roots, tuber and stems xi. fragmentation e.g. Spirogyra. b. Artificial vegetative propagation. i. budding ii. layering iii. grafting iv. cutings.	2. relate the types of reproduction to the success of the organisms. 3. assess the merits and demerits of each type of reproduction.	2. Do practical work on artificial vegetative propagation.	
3. Sexual reproduction in flowering plants. a. meaning - union of gametes b. meiosis in relation to gamete formation. c. Structure and function of the floral parts. i.e. bracts, receptacle, sepals petals, stamens, filaments,	Pupils should be able to: 1. assess the importance of sexual reproduction in plants. 2. to draw and label the various parts of the flower	1. Make longitudinal sections of a flower and draw its parts. 2. Examine different types of flowers and describe them.	- do -

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
anthers, ovule, ovary placenta, style, stigma and nectaries.	3. discuss the functions of different floral parts.		
d. Terms used to describe flowers i.e. superior, inferior, half inferior ovaries, unisexual flowers, monoecious and dioecious plants, bisexual etc.	4. describe different types of flowers using the correct terminologies.		
4. Pollination. a. meaning	5.define correctly the term pollination.	3. View a movie film on the process of pollination.	
b. Agents of pollination i. wind pollinated flowers ii. characteristics of wind pollinated flowers.	6.discuss different agents of pollination.	4. Examine flowers that are typically wind pollinated.	
iii. Insects pollinated flowers. iv. Characteristics of insect pollinated flowers.	7.relate the structure of a flower to its mode of pollination.		
c. Types of pollination i.self-pollination and devices which aid it.	8.discuss the merits and demerits of different methods of pollination.	5. To draw and label the parts of wind and insect pollinated flowers.	

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
ii. Cross-pollination and devices which aid it. d. Production of gametes-meiosis.			
5. Fertilization a. germination of pollen grain b. fusion of male gamete with a female gamete.	Pupils should be able to describe the process of fertilization in a flowering plant.	View a movie film on the process of pollination and fertilization.	- do -
6. Seed and fruit formation a. Types of seeds - Endospermous -Non endos-permous b. Type of fruits - simple, aggregate compound. - fleshy, dry dehiscent, dry indehiscent.	Pupils should be able to: 1. match the parts of the flower with the parts of the fruit or seed. 2. distinguish between endospermous and non-endospermous seeds. 3. distinguish different types of fruits.	1. Pupils to observe, dissect and draw different types of seeds and fruits eg. oranges, tomatoes, coconut, pear, bean, maize, cashewnut, groundnut. 2. Pupils to label the different parts of the seed and fruit.	- do -
7. Dispersal of seeds and fruits. a. Wind dispersal b. Water dispersal c. Animal dispersal d. Explosive mechanism e. Dispersal by man.	Pupils should be able to: 1. relate the structure of the seed/fruit with the methods of dispersal. 2. discuss the different methods of seed/fruit dispersal.		- do -

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
9. Sexual reproduction in higher animals.	Pupils should be able to: 1. relate the structure of the reproductive organs to the type of fertilization.	Make a dissection of fish, chicken and a mouse and locate the reproductive organs.	- do -
a. Sexual reproduction in fish. i. reproductive organs ii. spawning iii. Fertilisation iv. Incubation period v. Breeding habits vi. Parental care	2. distinguish internal fertilization from external fertilization.		
b. Sexual reproduction in birds i. Reproductive organs ii. Fertilization iii. Laying of eggs iv. Incubation period v. Breeding habits vi. Parental care.	3. discuss how the embryo is nourished during the incubation/gestation period. 4. discuss parental care in fish, birds and mammals		
c. Sexual reproduction in Man or mouse. i. reproductive organs ii. copulation and fertilization iii. gestation period iv. Parental care			

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
13. GENETICS	Pupils should be able to:		
a. Introduction - concept of heredity.	1. state correctly the meaning of heredity.	1. Construction of DNA molecule.	
b. Chromosomes and the theory of gene action (Transmission of characters).	2. describe how a character is transmitted from the parent to the offspring.	2. Drossophila breeding experiments.	
c. Nucleic acids structure of DNA and RNA, functions of DNA and RNA.	3. relate the structure of the DNA molecule to transmission of genetic factors and mitosis or meiosis.	3. Grow and breed garden peas of different characteristics.	- do -
d. Cell division mitosis and meiosis, DNA Duplication.	4. describe the effect of crossing two parents with one contrasting character.	4. Demonstration and determination of blood groups of all pupils in a class.	
e. Principles of Mendelian inheritance.	5. state correctly the meaning of Dominance and recessive.		
f. Hereditary factors: i. Meaning of dominance recessive, pure line, phenotype, genotype allele homozygous and heterozygous. ii. Mendel's first law of inheritance. iii. Partial dominance.			

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
g. Principles of plant and animal breeding (high production, resistance to infection) h. Blood groups. i) Maternal incompatibility ii) Blood transfusion.			
14. EVOLUTION 1. Meaning of evolution 2. Factors controlling evolution e.g. Food, space, Competition, mutation and recombination, genetic drift, change of environment. 3. Evidences of evolution-only an outline. 4. Natural Selection a. Meaning. b. How natural selection works. c. Variations and adaptations.	Pupils should be able to: 1. State correctly the meaning of the word evolution. 2. outline the forces controlling evolution. 3. discuss the role played by each factor in evolution. 4. describe evidences available to prove that evolution has taken place and is still taking place. 5. describe the mechanisms of natural selection.		- do -
15. INTERDEPENDENCE OF ORGANISMS (INTERELATIONSHIP OF ORGANISMS) 1. The concept of ecosystem a. Plant as producers	Pupils should be able to: 1. describe the relationship		

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
b. Animals as Consumers c. Microorganisms as decomposers. 2. Food chains and food webs (how food control the ecosystem). 3. Balance of nature (stabilasation of organisms in natural habitats) controlling factors other than food e.g. physical, chemical and biotic factors.	between producers, consumers and decomposers in energy flow. 2.discuss the role of each trophic level in the ecosystem. 3.discuss the factors controlling the stabilisation of organisms.		- do -
5. Insects as pests of crops a. To study the life cycle of maize stalk borer, cotton stainer coffee mealybug, and weevils. b. Discuss the effect of the above pests to crops (damages they cause) c. Control methods.	Pupils should be able to: 1. define the word pest. 2. describe the life cycle of insect pests studied. 3. discuss the harm caused by insect pests. 4. evaluate the control measures in relation to the life cycle of the pest.	Collect and observe damages caused by different pests.	- do -

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
ii. Parasites of animals A. Blood stream - Bacteria (VD) Protozoa (Plasmodium study the life cycle of Plasmodium and the life of its vector the mosquitoe)	3. relate the habitat and life cycle of the parasites to its mode of control.	3. Obtain blood samples from the hospital and study different stages of development in Plasmodium	
- Metazoa - Trypanosoma and Schistosoma (Study the life cycle of Trypanosoma and Schistosoma to elucidate the mode of infection).	4. describe the symptoms, transmission and control measures of the parasites studied.	4. Observe the external morphology of the adult and all stages in the life cycle of mosquitoe-Anopheles 5. Obtain samples of blood or permanent slides of Trypanosoma and Schistosoma to observe the external morphology.	- do -
NOTE: It is important for the teachers to emphasize the following. a. Methods of detecting the desease (symptoms) b. Methods of spread (Transmission) c. Control measures.		6. Where possible urine and stool samples could be examined (Teacher to demonstrate)	
B. Ectoparasites - meaning examples fleas, lice and ticks.		7. Examine specimens of fleas lice and ticks and also examine their habitats.	

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
16. PARASITES AND DISEASES OF MAN 1. Meaning and survival problems of parasite. a. What is a parasite? b. How does a parasite survive? c. The life cycles of Tapeworm, Roundworm/ Hookworm - to elucidate the adaptations to parasitic mode of life	Pupils should be able to explain the relationship between a parasite and its host.	Obtain and observe the different stages in the life cycle of these parasites e.g. eggs, 'Larvae' and adults. NB.consult the nearest hospital for help.	- do -
2. Infections of biological origin. Principles concerning the relationship between man and pathogenic agents. a. How human life facilitates infection for example. i.social (ie.Kissing-bacterial, contact-leprosy, burial ceremony - cholera) a. where they are found b. what are their effects	Pupils should be able to describe different methods of transmitting diseases.		

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
to their host c. control measures.			
5. Major endemic diseases in Tanzania a. Meaning of endemicity i. Localized (regional) e.g. Malaria, Bilhazia. ii. non localized (National eg.T.B., Dysentry, V.D.) b. Factors that govern endemicity e.g. climate, topography with regards to the disease studied i.e. Malaria, Bilhazia, Typhoid, V.D., Dysentry and Leprosy.	1. Pupils should be able to state correctly the meaning of endemicity. 2. Pupils should be able to distinguish localized from non localized endemicity, 3. Pupils should be able to relate the endemic diseases to climatic and topographic conditions in Tanzania.	Pupils to draw the map of Tanzania and Map out the distribution of endemic diseases.	- do -
17. MAN AND HIS NATURAL RESOURCES. 1. Meaning of Natural resources. 2. Types of Natural resources. A.Solar energy i.Natural uses ii.Artificial uses. B.Land i.Vegetation (Trees and grasses)	Pupils should be able to: 1. state correctly the meaning of natural resources. 2. outline the types of natural resources. 3. discuss the uses of each natural resource .	1. Visit national parks or game reserves. 2. local excursion to observe land use and management.	- do -

TOPIC OUTLINE	OBJECTIVES	SUGGESTED PUPIL ACTIVITY	REFERENCES
-Uses of vegetation -Conservation of vegetation. ii. Wild life - uses of wild life - conservation of wild life. C. Water Uses of water - industrial use - domestic use ii. Conservation of water. D. Air - Use of air conservation of air.	4. discuss methods used to conserve the natural resources.		- do -
3. Pollution a. Meaning of pollution b. Effects of pollution to living organisms. (refer to water, air and household pollution)	Pupils should be able to 1. State correctly the meaning of the word pollution. 2. evaluate the hazards of pollution to human life. 3. discuss control measures against pollution	Local excursion to locate polluted areas.	- do -

APPENDIX IIIa

Candidate's Number ....

THE UNITED REPUBLIC OF TANZANIA

NATIONAL EXAMINATIONS COUNCIL
CERTIFICATE OF SECONDARY EDUCATION
EXAMINATION, 1983

033/1 BIOLOGY PAPER 1

(For both School and Private Candidates)

TIME: 3 Hours.

INSTRUCTIONS TO CANDIDATES

1. This paper consists of sections A,B and C.
2. Answer ONE question in section A, and all questions in both Sections B and C
3. All answers must be written in the spaces provided for each question. Unless otherwise stated, answers written elsewhere will NOT be marked.
4. Write your centre and index number as indicated on the top right hand corner of every page.
5. All diagrams and labels should be in lead pencil. Other writting should be in blue or black ink or ball point pen.
6. FAILURE TO FOLLOW INSTRUCTIONS WILL LEAD TO LOSS OF MARKS.

FOR EXAMINER'S USE ONLY	
QUESTION	SCORE
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
TOTAL	

This paper consists of 17 printed pages.

- 2 - Candidate's No.........

SECTION A

Answer ONE question from this section. Each question carries 20% of the total marks.

(1) 1. Describe the various factors which cause loss of soil fertility and the methods which might be used in soil conservation.

(2) 2. Write an essay on how the knowledge of Biology has contributed to the development of our country and mankind as a whole.

(3) 3. Describe the mechanism of breathing in man.

Note: The circumscribed numbers have been assigned to each item by the author and constitute the item numbers referred to in all discussions in the study.

 Candidate's No.........

SECTION B

Answer ALL questions in this section. Questions (i) - (x) are multiple choice items. Write down the letter of the most correct answer in the box space besides each question. The first one is done as an example.

4. The three components of blood necessary for blood clotting are:-
   A. Leucocytes, thrombocytes and fibrinogen.
   B. Thrombocytes, prothrombin and fibrinogen.
   C. Erythrocytes, leucocytes, and thrombocytes. [B]
   D. Erythrocytes, leucocytes and fibrinogen.

(4) (i) In any ecosytem, every food chain usually:
   A. Begins with producers and ends up with big predators like lions.
   B. Begins with herbivores and ends up with big predators like lions.
   C. Begins with producers and ends up with decomposers, mainly bacteria. [ ]
   D. Begins with herbivores and ends up with carnivores which outnumber the herbivores.

(5) (ii) Aristotle, the ancient Greek Philosopher believed that all the food made by plants comes from decayed animal and plant material absorbed by plant roots from the soil. Later in the 17th century a tree seedling was planted in 90Kg of soil and left to grow for five years being given water only. After this period the plant gained 74Kg. in weight while the soil lost only 56 grams. Using your biological knowledge, the most suitable conclusion is that:
   A. There was an error in recording the readings because loss in weight of the soil could not have been that small.
   B. Most of the raw materials for food manufacture came from water and atmospheric carbondioxide. [ ]
   C. The food material for food manufacture come from water only.
   D. The raw materials for food manufacture are not obtained from the soil nor from water.

(6) (iii) With eyes stationary, one can see all the print on an average sized page of a book but cannot read all the words at the same time. In order to read every word, he has to move his eyes from left to right along the lines of print. The reason for this is:

A. When the eyes are stationary, the images of some of the words are focussed on the blind spot which do not contain sensitive cells.

B. When one looks at the whole page, the light from some of the words is deflected by the humours, and thus insufficient light falls on the retina.

C. One moves the eyes from left to right when reading because of training (conditioned reflex).

D. The image of each word has to be focussed on the fovea which contains millions of light sensitive cells.

☐

(7) (iv) Although all organisms are alike in general characteristics, they differ in their detailed characteristics because of their:

A. names.

B. shapes.

C. specialization.

D. habitat.

☐

(8) (v) A root hair absorbs water until it can no longer absorb any more. What force prevents further entry of water into the root hairs?

A. Turgor pressure.

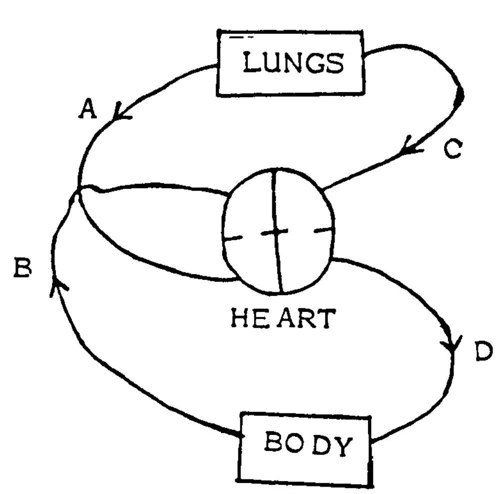
B. Osmotic pressure.

C. Suction pressure.

D. None of the above.

☐

(9) (vi) The figure below is a diagramatic representation of the circulation of blood in a mammal. The arrow heads show the direction of flow of blood.



Which of the following is correctly matched for a wrongly oriented blood vessel?

A. A - Pulmonary artery.

B. B - Vena cava.

C. C - Pulmonary vein.

D. D - Aorta.

(10) (vii) Which one of the following is a vector of sleeping sickness?

A. Housefly.

B. Mosquito.

C. Trypanosoma.

D. Tsetse fly.

(11) (viii) Keep in mind that a mature red blood cell is three dimensional. Suppose a needle should pierce it. Starting from the outside and moving to the centre, list in order the structures that would be pierced by the needle.

A. The cell membrane, cytoplasm nucleus.

B. The cell membrane, cytoplasm.

C. Cytoplasm, cell membrane, nucleus.

D. The cell wall, cell membrane, cytoplasm, nucleus.

(12) (ix) The most important outcome, of cellular respiration of living organisms is:

A. Discharge of carbondioxide.

B. Intake of oxygen.

C. Formation of water.

D. Release of energy.

(13) (x) The shortest pathway through which an impulse can travel from a stimulated sense organ to an effector organ is called:

A. Spinal reflex.

B. Reflex arc.

C. Reflex pathway.

D. Synapse.

...8/.

5. The following are matching items. Match the phrases or statements in list A with those in list B by choosing the correct number of words or phrase in list B and writing it against its relevant position in the space provided alongside list A.

List A

- (14) (i) __________ Coleoptile
- (15) (ii) __________ Causes rapid mobilization of glucose from the liver under conditions of stress.
- (16) (iii) __________ Birds and mammals differ from other vertebrates in having.
- (17) (iv) __________ Acts as a resevoir for blood cells of all kinds.
- (18) (v) __________ Insects.
- (19) (vi) __________ Autotrophic.
- (20) (vii) __________ Wings of birds and insects.
- (21) (viii) __________ Incomplete dominance.
- (22) (ix) __________ The basis of life.
- (23) (x) __________ Development of aphids from unfertilized eggs.

List B

1. Enclosing a radicle in members of the grass family.
2. Enclosing a plumule in members of the grass family.
3. Andrenaline.
4. Thyroxine.
5. The heart.
6. A double circulation.
7. The liver.
8. The Spleen.
9. Spiders and bees.
10. Praying mantis and termite.
11. Holophytic nutrition.
12. Holozoic nutrition.
13. Homologous structures.
14. Analogous structures.
15. Partial dominance.
16. Dominance recessiveness of two factors.
17. Cell cytoplasm.
18. Protoplasm.
19. Parthenocarpy.
20. Parthenogenesis.

6. The following statements are either true or false. In the spaces provided, write TRUE if the statement is true and FALSE if the statement is false. Number (i) has been done as an example.

TRUE. The chief characteristic of birds is the possession of feathers.

- (24) (i) ______________ Gametes are absolutely essential for reproduction.
- (25) (ii) ______________ Fossils are the remains of only dead mammals of the past preserved in rocks.
- (26) (iii) ______________ The human skeleton is made up of bones only.

(27) (iv) ____________ Among unicellular organisms the term 'growth' (defined in this case as an increase in the number of cells) is equivalent to the term reproduction.

(28) (v) ____________ In general the power of regeneration is greater in animals which are simple, while limited in more complex animals.

(29) (vi) ____________ Germ mother cells (those giving rise to gametes) have a diploid number of chromosomes in their nuclei.

(30) (vii) ____________ Oxygen liberated during photosynthesis comes from the Carbondioxide that is broken down during the overall process.

(31) (viii) ____________ When an animal is placed in an apparatus which ensures that it re-breathes the same air continuously, its breathing rate increases.

(32) (ix) ____________ Excretion and defecation mean the same thing.

(33) (x) ____________ In tropical areas some plants shed their leaves in order to reduce the rate of transpiration.

SECTION C

Answer ALL questions in this Section.

7. (34) (a) In a successful operation, a cow's alimentary tract was replaced by a hyena's alimentary tract. The cow recovered completely. If the cow continued to feed on the same diet as before it was operated, what problems would it face? Why?

____

____

____

____

(35) (b) Explain why it is possible to swallow food while you are upside down (standing on your head)?

____

____

____

(c) Substance X occurs in mammalian blood. When it occurs in the form in which it occurs in blood, the substance will turn brick red if Benedicts solution is added to it and heated. This substance can be converted to another form and be stored in the liver and muscles.

(36) (i) The name of substance X is ______________________

(37) (ii) The chemical formula of substance X is ______________________

(38) (iii) The form in which it is stored in the liver and muscles is ______________________

(39) (iv) The role it plays in the body is ______________________

____

____

____

(40) (v) Why is it called a monosaccharide?

____

____

8. (a) What role do the following play in the organisms in which they are found?

(41) (i) Mucus in the tracheal system of man.

____

____

(42) (ii) Ciliary muscles of the eyes.

____

____

____

____

(43) (iii) Auditory nerve. ______________________

____

____

____

(44) (iv) Haemoglobin. ______________________

____

____

____

(45) (v) Chlorophyll. ____________________

(b) Write short notes on the following:-

(46) (i) Oxygen debt. ____________________

(47) (ii) Sex cell. ____________________

(48) (iii) Dental formula. ____________________

(49) (iv) Sebaceous gland. ____________________

(50) (v) Vestigial structure. ____________________

(c) Differentiate between the following:-

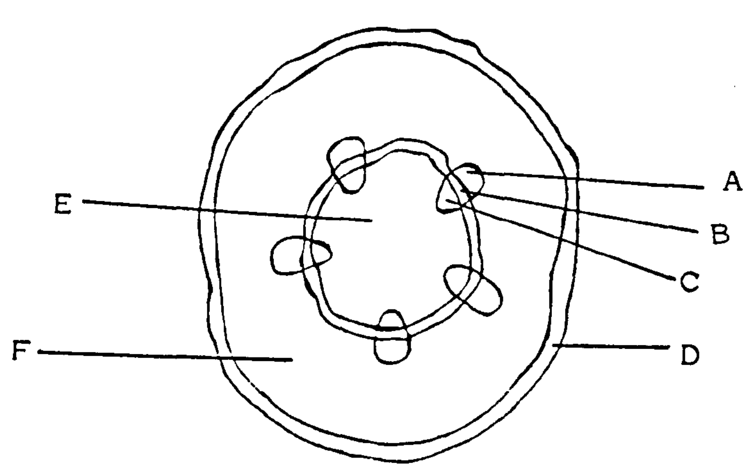
(51) (i) Urea and Urine. ____________________

(52) (ii) Vitreous humour and Aqueous humour.

(53) (iii) Food chain and food web.

(54) (iv) Larva and pupa.

9. (a) The diagram below is a T.S. of a part of a flowering plant.



(55) (i) Name the parts labelled A - F on the diagram.

A. ______________ B. ______________ C. ______________

D. ______________ E. ______________ F. ______________

(56) (ii) State the role of B and C.

The role of B ____

____

____

The role of C ____

____

____

(57) (iii) Giving reasons, state whether it is from a monocot or dicot plant. ____

____

____

____

(58) (iv) From what part of the plant was the T.S. made?

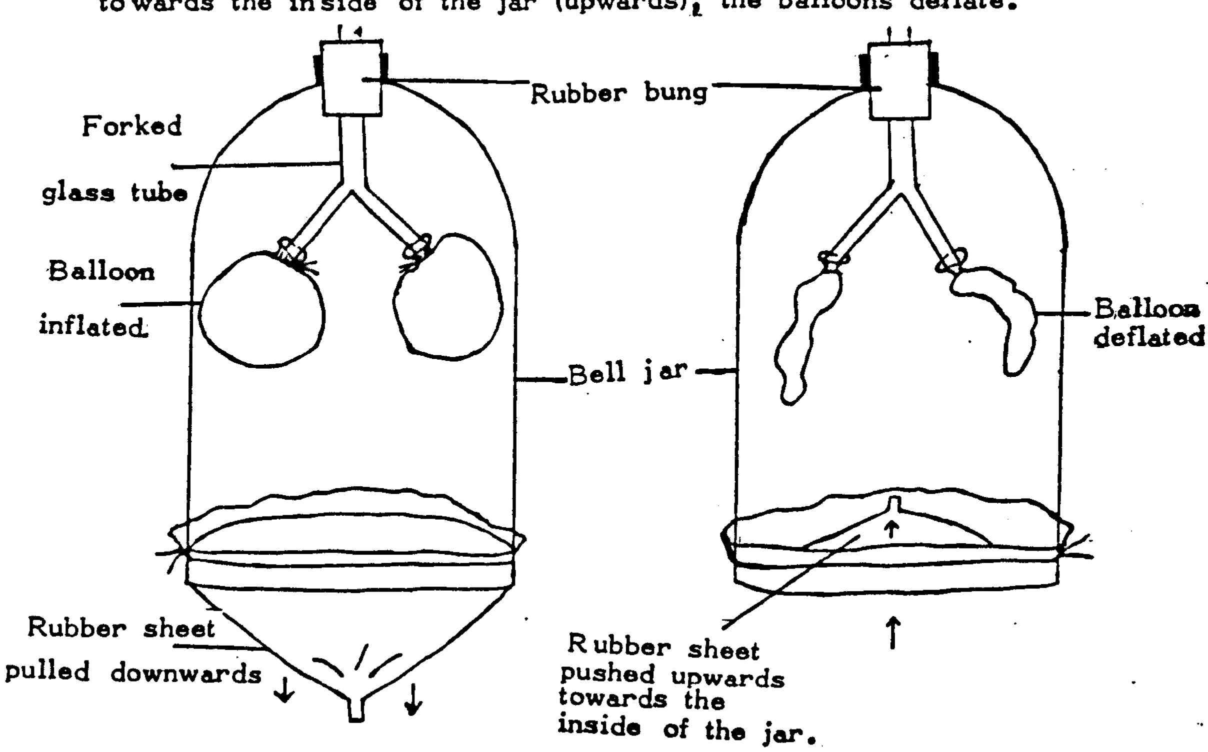
Give reasons. ____

____

____

____

(b) The diagram below is a model of the lungs and thorax. When the rubber sheet is pulled downwards the balloons inflate; when it is pushed towards the inside of the jar (upwards), the balloons deflate.



(59) (i) To what parts of a mammalian's body are the labelled parts of the apparatus equivalent?

---

(60) (ii) Which mechanism of the lung inflation is not illustrated by this apparatus?

---

(61) (iii) How is each labelled part of the appratus similar to, and yet different from the parts of a mammal's body to which it is equivalent.

(c) (i) Blood transfusions require the knowledge of the compatibility of the different blood groups. Sometimes the first transfusion may not bring complications but subsequent transfusions to the same recipient and from the same donor may bring complications. In the following table, put a tick (✓) where you think transfusions are not likely to bring complications, even in subsequent transfusions. Put a cross (x) where you think transfusions may bring complications.

(62)

		DONOR							
		A Rh(+)	A Rh(-)	B Rh(+)	B Rh(-)	AB Rh(+)	AB Rh(-)	0 Rh(+)	0 Rh(-)
RECIPIENT	A Rh(+)								
	AB Rh(-)								
	O Rh(+)								

(63) (ii) Giving reasons, explain why a group O person is not necessarily a universal donor, and when a group AB person can receive blood from any person irrespective of the blood group of the donor.

____

____

____

____

____

____

____

____

____

____

____

____

____

10. Explain the meaning of the following biological terms:-

(64) (i) Disease. ______________________________

(65) (ii) Parasite. ______________________________

(66) (iii) Tadpole. ______________________________

(67) (iv) Root hair. ______________________________

(68) (v) Trophic level. ______________________________

(69) (vi) Recessive gene ______________________________

(70) (vii) Soil. ______________________________

(71) (viii) Chordate. ______________________________

11. (a) (72) If the body temperature of a person suffering from fever rises above a certain degree the person may die. Explain why.

____

73 (b) By sitting completely still for several hours as can happen during a very long bus ride, many people experience the swelling of the lower parts of their legs. This condition quickly subsides when they begin to walk about. Explain.

____

____

____

____

____

____

74 (c) Explain why some trees are killed when a ring of bark is removed from their stems.

____

____

____

____

____

____

75 (d) How is a fertilized bird's egg similar to a bean seed?

____

____

____

____

- End -

APPENDIX IIIb

Candidate's No................

THE UNITED REPUBLIC OF TANZANIA

NATIONAL EXAMINATIONS COUNCIL
CERTIFICATE OF SECONDARY EDUCATION
EXAMINATION, 1983

033/2

BIOLOGY PAPER 2

(For School and Private Candidates)

Time: 2 Hours.

INSTRUCTIONS TO CANDIDATES

1. Answer all questions in this paper.
2. All answers must be written in the spaces provided for each question. Unless accompanied by tangible explanation from the supervisor, answers written elsewhere will NOT be marked.
3. All writing must be in blue or black ink, while all diagrams and their labels should be in lead pencil.
4. Write your centre and index number on the top right hand corner of every page.
5. Failure to adhere to instructions may lead to loss of marks.

FOR EXAMINER'S USE ONLY	
QUESTION	SCORE
1	
2	
3	
TOTAL	

This paper consists of 6 printed pages.

1. (i) (a) You have been provided with specimen A. Using reagents provided identify the food substances present in specimen A.

Write your procedure, observations and conclusions in the spaces provided:

PROCEDURE	OBSERVATION	INFERENCE

Candidate's No............

(2) (b) Name the end products of digestion of each food substance present in A.

Food Substance	Products

(3) (c) Name the digestive enzymes involved in the digestion of each of the food substances present in A.

Food Substance	Enzyme digesting it

(4) (d) State the medium in which the named enzymes act best. (Just indicate whether it is acidic or alkaline medium).

Enzyme	Medium

(5) (e) What age groups of man need most each food substance identified in 1.(a). Give reasons.

Food substance	Age group	Reasons

Candidate's No........

2. ⑥ (a) Identify specimen B and D.

Specimen B ______________

Specimen D ______________

⑦ (b) What role does the following play in the organism in which it is found?

Specimen B ______________

____

(c) You have been provided with Specimen D on which organisms are growing.

⑧ (i) The common name of the organisms growing on specimen D is:

____

⑨ (ii) The mode of nutrition of the organisms is:

____

⑩ (iii) How do the cells of these organisms differ from cells of flowering plants?

____

____

⑪ (iv) What role do these organisms play in the balance of nature?

____

____

____

 Candidate's No........

3. (a) (i) You are provided with specimen Y. Observe it carefully then draw and label its parts.

(12) Drawing of Specimen Y.

Specimen X is a peel of the lower side of specimen Y. Observe it carefully under the microscope.

(13) (ii) From your observations state the functions of specimen Y. Give reasons.

____

____

____

____

...6/.

(i) What is the common name of specimen M? (14)

____

(ii) What type of storage organ is specimen M? (15)

____

(iii) What makes it be referred to as a stem?

____ (16)

____

(iv) What role does it play in the life of the plant from which it is formed?

____ (17)

____

(v) Cut it into two halves then add a drop of iodine solution on the cut surface of one of the pieces. What do you observe?

____ (18)

____

(vi) From your observations what type of food substance does specimen M contain?

____ (19)

- End -

LIST OF SPECIMENS FOR THE PRACTICAL EXAMINATION

Specimen A	Egg albumen + sucrose in Solution
Specimen B	Mammalian kidney
Specimen D	Food material on which mould was growing
Specimen Y	Entire leaf (Hibiscus sp.)
Specimen X	Slide showing lower epidermis of Specimen Y
Specimen M	Round potato (Irish Potato)

APPENDIX IVa

INSTRUMENT I:FORM FOUR PUPILS

This is not an examination. It is a questionnaire in which are listed the biology topics and sub-topics which you have studied during the four years you have been at secondary school. Please give your opinion about the ease or difficulty of the topics by putting a (✓) in the appropriate box. When you have finished, you should have only 45 boxes ticked.

Easy	= I understood it first time I was taught it.
Average	= I had to work hard at it, but now I understand it.
Difficult	= I have worked hard at it, but I still do not understand it.
Not covered/Not taught	= I have not covered/I was not taught this topic.

TOPIC	SUBTOPIC	Easy	Average	Difficult	Not Covered
CLASSIFICATION	1 Characteristics of living things and classifications of things into living and non-living.				
	2 Principles for classifying plants into major groups. Seed plants - Angiosperms, Gymnosperms including monocots, dicots, pines and figs.				
	3 Non-seed plants - algae, fungi, mosses and ferns.				
	4 Principles for classifying animals into major groups of: Invertebrates Protozoans (Amoeba, Paramecium,Trypanosoma, Plasmodium) Metazoa; the worms (flat,round segmented) Arthropods (insects, crabs, spiders)				

TOPIC		SUBTOPIC				
	5	Vertebrates: fish, Amphibian Reptiles, Birds and Mammals.				
CELL STRUCTURE AND FUNCTION	6	Structure and function of cell parts.				
	7	Comparison of plant and animal cells.				
SOIL AND PLANT GROWTH	8	Soil formation. Constituents and properties of soils.				
	9	Soil and water (water retention, capillarity and porosity).				
	10	Soil fertility and plant growth including the Nitrogen cycle.				
NUTRITION	11	<u>Plant Nutrition</u>: Leaf structure and function. photosynthesis. Plant mineral nutrition. Synthesis of fats and proteins. Plant storage organs.				
	12	<u>Animal Nutrition</u>: Classes of food and food tests. Balanced diet (including Vitamins and deficiency diseases).				
	13	Enzymes and their characteristics. Digestion and digestive system of mammals. Absorption and assimilation.				

TOPIC		SUBTOPIC				
NUTRITION	14	Special modes of nutrition (symbiosis, saprophytism, parasitism)				
	15	Food preservation and storage (the need for and methods of).				
TRANSPORT SYSTEM	16	The relationship between size, surface area and the need for transport system in complex/large organisms.				
	17	Transport in plants: Structure of roots and stems. Water movement in plants (including osmosis, turgidity, wilting, plasmolysis, and diffusion). Transpiration and conditions affecting it (including transpiration pull and root pressure).				
	18	Transport in Mammals: Composition and function of mammalian blood. Structure of heart and blood vessels. Blood circulation. Lymph and lymphatic system.				
RESPIRATION	19	Gaseous exchange in Animals: Structure and function of respiratory organs (skin, tracheal system of insects, gills and lungs). Mechanism of gaseous exchange including Breathing and diffusion of gases.				

TOPIC		SUBTOPIC				
RESPIRATION (cont'd)	20	Tissue respiration: Oxidation of food to release energy. Release and utilisation of chemical energy (ATP and ADP) Anaerobic respiration (fermentation).				
	21	Gaseous exchange in plants: Structure and function of stomata and lenticels. Comparison of photosynthesis and respiration.				
	22	Effects of smoking on health.				
EXCRETION AND OSMO-REGULATION IN MAMMALS	23	Structure and function of kidneys, skin, lungs and liver.				
CONTROL AND CO=ORDINATION SYSTEM IN ORGANISMS	24	The Nervous system in mammals: Structure and function of nerve cells, brain and spinal cord. Cranial and spinal nerves.				
	25	Structure and function of the sensory organs. (Eye, ear, skin, tongue and nose.)				
	26	Endocrine System in mammals: The main endocrine glands, hormones and their functions.				
	27	Response to stimuli in green plants: Auxins and tropisms.				

TOPIC	SUBTOPIC				
MOVEMENT	28 Movements using cilia, flagella and amoeboid movements.				
	29 Skeleton and joints: Structure and function of endoskeleton and exoskeleton.				
	30 Types, structure and arrangement of muscles (smooth, striated, cardiac)				
	31 Movements in vertebrates (fish, frogs, snakes, mammals).				
GROWTH AND DEVELOPMENT	32 The process of cell division (mitosis), enlargement and differentiation.				
	33 Seed structure and function: Germination and conditions necessary for germination to occur. Primary and secondary growth in plants.				
	34 Growth in insects (complete and incomplete metamorphosis), amphibians, birds, reptiles and mammals.				
REPRODUCTION	35 Sexual reproduction, meiosis, and gamete formation.				

TOPIC		SUBTOPIC	Easy	Average	Difficult	Not Covered
REPRODUCTION	36	Structure and function of floral parts. Agents and types of pollination. The process of fertilisation. Seed and fruit formation and dispersal. Asexual reproduction and vegetative propagation.				
	37	Sexual reproduction in mammals.				
GENETICS	38	Chromosomes, genes, mutation. Structure and function of DNA. Mitosis and meiosis.				
	39	Mendel's first law of inheritance (dominant and recessive genes, phenotype and genotype).				
	40	The importance of genetics in plant and animal breeding.				
EVOLUTION	41	Evidences of evolution. The process of change (factors controlling evolution). The process of natural selection.				
ECOLOGY AND INTERDEPENDENCE OF ORGANISMS	42	Ecosystems Population and communities succession. Factors affecting distribution of organisms. Food chains and food webs.				

TOPIC	SUBTOPIC	Easy	Average	Difficult	Not Covered
PARASITES & DISEASES OF MAN	Man and pathogens: 43 Different types of pathogens and disease causes.				
	Major endemic diseases in Tanzania (Bilharzia, typhoid, malaria, leprosy, T.B., dysentry). Causes, methods of transmission and control.				
MAN & NATURAL RESOURCES	Conservation of the natural resources. 45 Pollution of environment (water, air and land pollution).				

Did you find it difficult to follow the instructions given?

Yes______________.

No______________.

APPENDIX IVb.

INSTRUMENT II: BIOLOGY TEACHERS.

Introduction:

We are trying to evaluate the O-level Biology Curriculum by considering the curriculum topics, curriculum objectives, and the O-level examination using the O-level Syllabus (1976), and the Teachers' Guide (1978).

This evaluation is designed to identify any problems which may affect the teaching and learning of O-level biology. The information obtained may also be useful for planning in-service-courses and seminars for biology teachers.

In view of this, you are kindly requested to respond to the following questionnaire which has five sections.

Your answers will be treated as strictly confidential throughout the course of this work.

Section I: The School and the Biology Teacher.

1. Name of school.________________________________

Please tick the following:

2. Location of school: Urban____________________

Rural____________________

3. Type of students attending the school:

Boys only____________________

Girls only____________________

Co-ed ____________________

4. Type of school: Boarding ____________________

Day ____________________

Day/Boarding ____________________

5. Sex of teacher Male ____________________

Female ____________________

6. Number of years teaching biology:

a) At present school ____________________yrs.

b) At other schools ____________________yrs.

c) Total ____________________yrs.

Section II: The Curriculum Topic:

Below is a summary of the biology topics as outlined in the Syllabus (1976) and elaborated in the Teachers' Guide (1978), which are to be covered within the four years of O-level secondary education.

Please indicate the relative ease or difficulty students have in understanding these topics by placing a (√) in the appropriate box using the following key:

Easy : Students understood it first time I taught it.
Average : I had to work hard at it, but now they understand it.
Difficult : I have worked hard at it, but still they don't understand it.
Not covered/not taught: I have not covered/not taught this topic.

TOPIC	SUBTOPIC	Easy	Average	Difficult	Not Covered
CLASSI-FICATION	1 Characteristics of living things and classifications of things into living and non-living.				
	2 Principles for classifying plants into major groups. Seed plants - Angiosperms, Gymnosperms including mono-cots, dicots, pines and figs.				
	3 Non-seed plants - algae, fungi, mosses and ferns.				
	4 Principles for classifying animals into major groups of: Invertebrates Protozoans (Amoeba, Paramecium Trypanosoma, Plasmodium) Metazoa; the worms (flat, round segmented) Arthropods (insects, crabs, spiders)				

TOPIC	SUBTOPIC				
	5 Vertebrates: fish, Amphibian Reptiles, Birds and Mammals.				
CELL STRUCTURE AND FUNCTION	6 Structure and function of cell parts.				
	7 Comparison of plant and animal cells.				
SOIL AND PLANT GROWTH	8 Soil formation. Constituents and properties of soils.				
	9 Soil and water (water retention, capillarity and porosity).				
	10 Soil fertility and plant growth including the Nitrogen cycle.				
NUTRITION	11 Plant Nutrition: Leaf structure and function. photosynthesis. Plant mineral nutrition. Synthesis of fats and proteins. Plant storage organs.				
	12 Animal Nutrition: Classes of food and food tests. Balanced diet (including Vitamins and deficiency diseases).				
	13 Enzymes and their characteristics. Digestion and digestive system of mammals. Absorption and assimilation.				

TOPIC		SUBTOPIC				
NUTRITION	14	Special modes of nutrition (symbiosis, saprophytism, parasitism)				
	15	Food preservation and storage (the need for and methods of).				
TRANSPORT SYSTEM	16	The relationship between size, surface area and the need for transport system in complex/large organisms.				
	17	Transport in plants: Structure of roots and stems. Water movement in plants (including osmosis, turgidity, wilting, plasmolysis, and diffusion). Transpiration and conditions affecting it (including transpiration pull and root pressure).				
	18	Transport in Mammals: Composition and function of mammalian blood. Structure of heart and blood vessels. Blood circulation. Lymph and lymphatic system.				
RESPIRATION	19	Gaseous exchange in Animals: Structure and function of respiratory organs (skin, tracheal system of insects, gills and lungs). Mechanism of gaseous exchange including Breathing and diffusion of gases.				

TOPIC		SUBTOPIC				
RESPIRATION (cont'd)	20	Tissue respiration: Oxidation of food to release energy. Release and utilisation of chemical energy (ATP and ADP) Anaerobic respiration (fermentation).				
	21	Gaseous exchange in plants: Structure and function of stomata and lenticels. Comparison of photosynthesis and respiration.				
	22	Effects of smoking on health.				
EXCRETION AND OSMO-REGULATION IN MAMMALS	23	Structure and function of kidneys, skin, lungs and liver.				
CONTROL AND CO=ORDINATION SYSTEM IN ORGANISMS	24	The Nervous system in mammals: Structure and function of nerve cells, brain and spinal cord. Cranial and spinal nerves.				
	25	Structure and function of the sensory organs. (Eye, ear, skin, tongue and nose.)				
	26	Endocrine System in mammals: The main endocrine glands, hormones and their functions.				
	27	Response to stimuli in green plants: Auxins and tropisms.				

TOPIC		SUBTOPIC				
MOVEMENT	28	Movements using cilia, flagella and amoeboid movements.				
	29	Skeleton and joints: Structure and function of endoskeleton and exoskeleton.				
	30	Types, structure and arrangement of muscles (smooth, striated, cardiac)				
	31	Movements in vertebrates (fish, frogs, snakes, mammals).				
GROWTH AND DEVELOPMENT	32	The process of cell division (mitosis), enlargement and differentiation.				
	33	Seed structure and function: Germination and conditions necessary for germination to occur. Primary and secondary growth in plants.				
	34	Growth in insects (complete and incomplete metamorphosis), amphibians birds, reptiles and mammals.				
REPRODUCTION	35	Sexual reproduction, meiosis, and gamete formation.				

TOPIC	SUBTOPIC	Easy	Average	Difficult	Not Covered
REPRO- DUCTION	36 Structure and function of floral parts. Agents and types of pollination. The process of fertilisation. Seed and fruit formation and dispersal. Asexual reproduction and vegetative propagation.				
	37 Sexual reproduction in mammals.				
GENETICS	38 Chromosomes, genes, mutation. Structure and function of DNA. Mitosis and meiosis.				
	39 Mendel's first law of inheritance (dominant and recessive genes, phenotype and genotype).				
	40 The importance of genetics in plant and animal breeding.				
EVOLUTION	41 Evidences of evolution. The process of change (factors controlling evolution). The process of natural selection.				
ECOLOGY AND INTERDEPEND- ENCE OF ORGANISMS	42 Ecosystems Population and communities succession. Factors affecting distribution of organisms. Food chains and food webs.				

TOPIC	SUBTOPIC	Easy	Average	Difficult	Not Covered
PARASITES & DISEASES OF MAN	43 Man and pathogens: Different types of pathogens and disease causes.				
	Major endemic diseases in Tanzania (Bilharzia, typhoid, malaria, leprosy, T.B., dysentry). Causes, methods of transmission and control.				
MAN & NATURAL RESOURCES	45 Conservation of the natural resources. Pollution of environment (water, air and land pollution).				

Section III: Curriculum Objectives:

Below is a list of objectives prepared by summarizing those listed in the Teachers' Guide (1978). They indicate the skills that students should possess at the end of their four-year biology course.

How important do you consider these objectives to be in an O-level biology course?

Please indicate your view by placing a (✓) in the appropriate box using the following key:

1. = of no importance.
2. = of some importance.
3. = very important.

	Objective	1	2	3
1.	Ability to: Show knowledge and understanding of biological facts, concepts, principles and theories.			
2.	Comprehend biological texts.			
3.	Apply biological principles to real life experiences.			
4.	Plan experiments.			
5.	Follow instructions in carrying out experimental and other biological procedures.			
6.	Make and record observations accurately.			
7.	Interpret data to draw conclusions and make inferences.			
8.	Draw biological diagrams accurately.			
9.	Analyse and solve problems.			
10.	Formulate hypotheses from numerical data, graphs, tables and other evidences.			
11.	Explain and discuss.			
12.	Write accurate reports of experiments and other biological procedures.			
13.	Use manual skills to set up and operate laboratory apparatus.			

Please state any other curriculum objectives which you feel are important in the biological education of pupils at this level and which have not been included in the above list:-

____

____

____

____

____

Section IV: Examination Objectives:

Below is the same list of objectives. Please now indicate by ranking them in order of importance from 1 - 13 those abilities that an O-level examination should test.

	Objective	Rank order
1.	Ability to: Show knowledge and understanding of biological facts, concepts, principles and theories.	
2.	Comprehend biological texts.	
3.	Apply biological principles to real life experiences.	
4.	Plan experiments.	
5.	Follow instructions in carrying out experimental and other biological procedures.	
6.	Make and record observations accurately.	
7.	Interpret data to draw conclusions and make inferences.	
8.	Draw biological diagrams accurately.	
9.	Analyse and solve problems.	
10.	Formulate hypotheses from numerical data, graphs, tables and other evidences.	
11.	Explain and discuss.	
12.	Write accurate reports of experiments and other biological procedures.	
13.	Use manual skills to set up and operate laboratory apparatus.	

Please list any objectives from the above which you feel are not currently examined in the Tanzanian O-level biology examination.

____

____

____

____

____

Section V: Practical Biology Examination:

1. How many times have you been involved in the preparation of the O-level biology practical examination?

   1. At present school ________ years.
   2. At other schools ________ years.

2. For the last practical examination (1982), indicate

   1. No. of students ___________.
   2. No. of Labs. in the school. ___________.
   3. Capacity of Labs, on a teaching day. ___________.
   4. Capacity of Labs. on the examination day. ___________.

3. Do you usually manage to do the practical examination in a single sitting?

   1. Yes______.
   2. No ______.

   If NO how many sittings did you have last time (1982)?

   No. of sittings last year________.

4. To what extent do the following issues pose problems during:
   (a) Preparation of the examination.
   (b) The actual day of examination.

	ISSUE	No Problem 1	Minor 2	Major 3
1.	Collection of Specimens			
2.	Identification of Specimens			
3.	Availability of time			
4.	Availability of equipment			
5.	Preparation of reagents			
6.	Safe storage of specimens			
7.	Co-operation from the Head of School			
8.	Co-operation from external examiner			
9.	Co-operation from candidates			
10.	Availability of laboratory technician			
11.	Size of laboratory			
12.	Organization of candidates			
13.	Co-operation from Examination Council			

Detail other problems experienced:

____

____

____

5. Is there a time when you were asked to collect a specimen with which you were not familiar?

Yes__________.

No__________.

If YES. What was the specimen?

Animal specimen__________.

Plant specimen __________.

Both __________.

6. Who helped you to identify the correct specimen?

Examination Council __________.

Fellow teacher at same school __________.

Biology teacher of another school __________.

Other person (please identify) __________.

7. How would you rate the practical examination of 1982?

Too easy __________.

Appropriate standard __________.

Too difficult __________

8. Was the time allocated for the practical examination last year (1982):

Too little? __________.

Appropriate? __________.

Too much? __________.

9. Again, below is the list of objectives prepared by summarizing those listed in the Teachers' Guide (1978). They indicate the skills that students should possess at the end of their four years biology course. Please indicate by putting (√) against any of those abilities which you feel can <u>best</u> be examined through <u>a practical examination</u>.

	Abilities	By Practical Examination
1.	Ability to: Show knowledge and understanding of biological facts, concepts, principles and theories.	
2.	Comprehend biological texts.	
3.	Apply biological principles to real life experiences.	
4.	Plan experiments.	
5.	Follow instructions in carrying out experimental and other biological procedures.	
6.	Make and record observations accurately.	
7.	Interpret data to draw conclusions and make inferences.	
8.	Draw biological diagrams accurately.	
9.	Analyse and solve problems.	
10.	Formulate hypotheses from numerical data, graphs, tables and other evidences.	
11.	Explain and discuss.	
12.	Write accurate reports of experiments and other biological procedures.	
13.	Use manual skills to set up and operate laboratory apparatus.	

10. What time of the year do you start examination revision exercises in practical biology?

____

11. What are your general views of the fairness of the method of conducting practical examination?

____

____

____

____

12. a) What would you prefer?

Practical examinations to continue as they are ______________

Practical examinations to be abolished but practical skills to be assessed through continuous assessment ______________

b) If practical examinations are to continue, are there any suggestions you would make for improvement?

i) ____

ii) ____

iii) ____

iv) ____

c) What would be the main disadvantage, if any, if practical examinations were to be abolished?

i) ____

ii) ____

iii) ____

iv) ____

Section Six: Continuous Assessment:

1. How often do you make use of the continuous assessment guidelines from the Examination Council?

very often ____________.
rarely ____________.
I never use it____________.

If you never use it, what other criteria, other than the guidelines do you use? Please describe.

____
____
____
____

How would you describe the guidance given in the guidelines for carrying out continuous assessment in biology?

too general ____________.
just adequate____________.
too detailed ____________.

If too general, what other information would you prefer added.

____
____
____
____

If to detailed, what details would you prefer removed.

____
____
____
____

2. Below is a list of methods which could be used for obtaining a continuous assessment mark for biology. Please indicate the order of your use of the methods by ranking them from 1 (most used) to 6 (least used).

	Method of assessment	rank order
1.	classwork exercise of ongoing work	
2.	class tests specially prepared	
3.	practical exercise of ongoing work	
4.	practical test specially prepared	
5.	term examination	
6.	project work or special studies	

3. To what extent do the following issues pose problems in carrying out continuous assessment? Tick (✓) in the appropriate column.

	Issues	Major	Minor	No Problem
1.	Availability of time			
2.	Preparation of exercise			
3.	Preparation of tests			
4.	Marking exercises/tests			
5.	Entering marks on CA record forms			
6.	Co-operation from Head of School			
7.	Student Absenteeism			
8.	Co-operation from Examination Council			
9.	Availibility of CA record forms			
10.	Co-operation from students			
11.	Teacher/pupil relationship			

Please detail other problems experienced.

____

____

____

APPENDIX IVc

INSTRUMENT III: BIOLOGY PANEL MEMBERS.

Introduction:

We are trying to evaluate the O-level Biology Curriculum by considering the curriculum topics, curriculum objectives, and the O-level examination using the O-level Biology Syllabus (1976), and the Teachers' Guide (1978).

This evaluation is designed to identify any problems which may affect the teaching and learning of O-level biology. The information obtained may also be useful for planning in-service-courses and seminars for biology teachers.

In view of this, you are kindly requested to respond to the following questionnaire which has three sections.

Your answers will be treated as strictly confidential throughout the course of this work.

Section I: Curriculum Objectives:

Below is a list of objectives prepared by summarizing those listed in the Teachers' Guide (1978). They indicate the skills that students should possess at the end of their four-year biology course.

How important do you consider these objectives to be in an O-level biology course?

Please indicate your view by placing a (√) in the appropriate box using the following key:

1. = of no importance.
2. = of some importance.
3. = very important.

	Objective	1	2	3
1.	Ability to: Show knowledge and understanding of biological facts, concepts, principles and theories.			
2.	Comprehend biological texts.			
3.	Apply biological principles to real life experiences.			
4.	Plan experiments.			
5.	Follow instructions in carrying out experimental and other biological procedures.			
6.	Make and record observations accurately.			
7.	Interpret data to draw conclusions and make inferences.			
8.	Draw biological diagrams accurately.			
9.	Analyse and solve problems.			
10.	Formulate hypotheses from numerical data, graphs, tables and other evidences.			
11.	Explain and discuss.			
12.	Write accurate reports of experiments and other biological procedures.			
13.	Use manual skills to set up and operate laboratory apparatus.			

Do you consider the above list to constitute a valid summary of those objectives outlines in the Teachers' Guide (1978)?

Yes ___________

No ___________

If no, please give your reasons e.g. by stating any other objectives which you feel are important in the biological education of pupils at this level and which have not been included in the above list.

____

____

____

____

____

____

Section II: Examination Objectives:

Below is the same list of objectives. Please now indicate by ranking them in order of importance from 1 - 13 those abilities that an O-level examination should test.

	Objective	rank order
1.	Ability to: Show knowledge and understanding of biological facts, concepts, principles and theories.	
2.	Comprehend biological texts.	
3.	Apply biological principles to real life experiences.	
4.	Plan experiments.	
5.	Follow instructions in carrying out experimental and other biological procedures.	
6.	Make and record observations accurately.	
7.	Interpret data to draw conclusions and make inferences.	
8.	Draw biological diagrams accurately.	
9.	Analyse and solve problems.	
10.	Formulate hypotheses from numerical data, graphs, tables and other evidences.	
11.	Explain and discuss.	
12.	Write accurate reports of experiments and other biological procedures.	
13.	Use manual skills to set up and operate laboratory apparatus.	

Please list any objectives from the above which you feel are not currently examined or are not given adequate coverage in the Tanzanian O-level biology examination.

____

____

____

____

____

Section III: Practical Examination.

Again, below is the list of objectives prepared by summarizing those listed in the Teachers' Guide (1978). They indicate the skills that students should possess at the end of their four years biology course. Please indicate by putting a (✓) against any of those abilities which you feel can best be examined through a practical examination.

	Abilities	By Practical Examination
1.	Ability to: Show knowledge and understanding of biological facts, concepts, principles and theories.	
2.	Comprehend biological texts.	
3.	Apply biological principles to real life experiences.	
4.	Plan experiments.	
5.	Follow instructions in carrying out experimental and other biological procedures.	
6.	Make and record observations accurately.	
7.	Interpret data to draw conclusions and make inferences.	
8.	Draw biological diagrams accurately.	
9.	Analyse and solve problems.	
10.	Formulate hypotheses from numerical data, graphs, tables and other evidence.	
11.	Explain and discuss.	
12.	Write accurate reports of experiments and other biological procedures.	
13.	Use manual skills to set up and operate laboratory apparatus.	

1. What are your general views of the fairness of the method of conducting the practical examination?

____

____

____

____

2. a) What would you prefer?

   Practical examinations to continue as they are ________________

   Practical examinations to be abolished but practical skills to be assessed through continuous assessment ________________

   b) If practical examinations are to continue, are there any suggestions you would make for improvement?

   i) ____

   ii) ____

   iii) ____

   iv) ____

   c) What would be the main disadvantages, if any, if practical examinations were to be abolished?

   i) ____

   ii) ____

   iii) ____

   iv) ____

APPENDIX IVd

INSTRUMENT IV: BIOLOGY EXAMINERS.

Introduction:

An attempt is being made to evaluate the Tanzanian O - level Biology Curriculum by considering the curriculum topics, curriculum objectives, and the O - level examination using the O - level Biology Syllabus (1976), the Teacher's Guide (1978), and the 1983 O -level biology examination. (Certificate of Secondary Education Examination).

This evaluation is designed to identify any problems which may affect the teaching and the learning of O - level biology. The information obtained may be useful for planning in-service-courses and seminars for biology teachers.

Part of the evaluation involves a consideration of the objectives, and the topic areas being examined.

In view of this, you are kindly requested to classify the 1983 examination questions according to objectives, and topic areas being examined, as indicated overleaf. Please, classify Paper I and Paper II on separate copies of Analysis Form I and II.

MATERIALS:

1. Analysis Form I: This consists of a list of the objectives prepared by summarizing those listed in the Teacher's Guide (1978).

2. Analysis Form II: This consists of a summary of the biology topics as outlined in the Syllabus (1976) and elaborated in the Teacher's Guide (1978). (The elaborated list of topic areas is also enclosed).

3. Analysis Form III: This consists of a grid matrix on which are listed the 13 objectives on the vertical axis, and Bloom's categories of educational objectives on the horizontal axis.

4. A short description of Bloom's Six Categories of Educational Objectives in the Cognitive Domain with examples of statements of instructional objectives. (See Appendix IVe).

5. The 1983 examination question papers.

TASK

1. Please, carefully study the examination papers for 1983.

2. Using the analysis Form I, classify each examination question according to the objective which you feel the question is examining. Each item (Section B and C) or complete question (Section A) may test either a single objective or several objectives.
When you feel several objectives are being tested, please enter the item or question number in the grid alongside each objective tested.

3. Using the analysis Form II, (and the elaborated list of topic areas, if need be), classify each examination question or item according to the topic area or areas which you feel are being examined.

4. Using analysis Form III, please classify the 13 objectives, as far as possible, according to Bloom's six categories of educational objectives by placing a (✓) in the appropriate box. A short description of Bloom's six categories of the Cognitive Domain is enclosed. Omit any of the objectives listed 1 - 13, which, in your view, do not fit into the taxonomy.

ANALYSIS FORM I.

Classification of the questions.

Year 1983 . Paper ________ .

	Objective	Number or letter of examination question or item in which the objective is tested.
	Ability to:	
1.	Show knowledge and understanding of biological facts. concepts, principles and theories.	
2.	Comprehend biological texts.	
3.	Apply biological principles to real life experiences.	
4.	Plan experiments.	
5.	Follow instructions in carrying out experimental and other biological procedures.	
6.	Make and record observations accurately.	
7.	Interpret data to draw conclusions and make inferences.	
8.	Draw biological diagrams accurately.	
9.	Analyse and solve problems.	
10.	Formulate hypotheses from numerical data, graphs, tables and other evidences.	
11.	Explain and discuss.	
12.	Write accurate reports of experiments and other biological procedures.	
13.	Use manual skills to set up and operate laboratory apparatus.	

FORM II.

Classification of Questions:

Year 1983 Paper _____.

	Curriculum Topic	Number or letter of examination question or item in which the topic is tested.
1	Classification	
2	Cell structure and function	
3	Soil and plant growth	
4	Nutrition in plants and animals	
5	Respiration in plants and animals	
6	Transport system in plants and animals	
7	Excretion and Osmoregulation in animals	
8	Control and co-ordination systems in organisms	
9	Movement	
10	Growth and Development	
11	Reproduction	
12	Genetics	
13	Evolution	
14	Ecology and interdependence of organisms	
15	Parasites and diseases of man	
16	Man and his natural resources	

ANALYSIS FORM III

CLASSIFICATION OF OBJECTIVES:

	Objective	Category of Educational Objective According to Bloom's Taxonomy of the Cognitive Domain.					
		Knowledge	Compre-hension	Appli-cation	Analysis	Synthesis	Evaluation
	Ability to:						
1	Show knowledge and understanding of biological facts, concepts, principles and theories.						
2.	Comprehend biological texts.						
3.	Apply biological principles to real life experiences.						
4.	Plan experiments.						
5.	Follow instructions in carrying out experimental and other biological procedures.						
6.	Make and record observations accurately.						
7.	Interpret data to draw conclusions and make inferences.						
8.	Draw biological diagrams accurately.						
9.	Analyse and solve problems.						
10.	Formulate hypotheses from numerical data, graphs, tables and other evidences.						
11.	Explain and discuss.						
12.	Write accurate reports of experiments and other biological procedures.						
13.	Use manual skills to set up and operate Laboratory apparatus.						

APPENDIX IVe

A short description of Bloom's Six Categories of Educational Objectives in the Cognitive Domain with examples of statements of instructional objectives.

Category	Description	Examples of statements of Instructional Objectives.
Knowledge	Involves the remembering of previously learnt information. It may involve the recall of specific facts, principles, theories, or concepts. At the knowledge level, the basic requirement is the bringing to mind of the appropriate information when called for.	Ability to show knowledge of:- common terms, - general and specific facts - methods and procedures - basic principles, concepts and theories.
Comprehension	Involves the ability to understand the meaning of learnt facts, principles, theories, concepts and any other information. When there is under-standing, one can translate information from one form to another; for example, numbers to words, to charts and give summary information when asked to.	Ability to show an understanding of facts, principles, concepts and theories. Ability to interpret information given in charts, graphs, and tables. Ability to justify methods and procedures used in experimental set ups.
Application	This category involves the ability to make use of learnt information in new situations. It may involve the application of learnt laws, principles, theories, concepts and procedures.	Ability to use methods and procedures in practical situations. Ability to interpret and convert data/information given in tables, charts and graphs from one form into the other. Ability to apply principles, concepts, theories and laws to new and/or practical situations. Ability to solve practical problems.

Analysis	Involves the ability to separate information by breaking it into its components so that it is easier to follow and therefore easier to understand. The basic requirement before making an analysis is the understanding of both the content and the structural format of the material or information.	Ability to recognize and/or make assumptions in support of information. Ability to give reasons for certain practical set-ups and components. Ability to reason and make inferences. Ability to differentiate and make comparisons. Ability to draw own charts, graphs, diagrams and make tables.
Synthesis	Involves the ability to plan and organise or re-organise information/material through creativeness.	Ability to plan experiments and other data collection procedures. Ability to write up reports of experiments, projects and other information needing logical sequence. Ability to formulate methods and procedures.
Evaluation	This is concerned with the ability to assess the value of information/material under specified conditions. In order to evaluate, one must know, understand, apply, analyse and synthesize data or any other information.	Ability to draw conclusions from given data and other information. Ability to interpret and summarize reports and experimental plans. Ability to criticize methods and procedures.

APPENDIX IVf

INSTRUMENT V: SOME BIOLOGY TEACHERS.

A semi-structured Interview Schedule for gathering additional information on Continuous Assessment in Biology.

1. What are the procedures used by the biology teachers to select exercises, tests and projects?

Lead question:

The continuous assessment record forms show places for recording marks for exercises, tests and projects done by students. How do you select the exercises, tests and projects?

Classwork exercise of ongoing work	
Class tests specially prepared	
Practical exercise of ongoing work	
Practical tests specially prepared	
Term examination	
Project work or special studies	
Homework	
Others:	

Space for additional comments:

2. What are the methods used by teachers to determine the criteria for assessment (i.e. to determine the skills and abilities to be assessed by the exercises, tests and projects)?

Lead question:

Is there a guide from the Examinations Council which tell you the abilities and skills to assess?

___

___

___

If not, how do you determine the abilities, skills to assess?

Mark is awarded if a student:

Show knowledge and understanding of facts, concepts, principles	
Plan experiments	
Make and record observations accurately	
Follow instructions in carrying out experiments and other procedures	
Draw biological diagrams accurately	
Explain and discuss	
Comprehend biological texts	
Apply biological principles to real life experiences	
Interpret data to draw conclusions and make inferences	
Write accurate reports of experiments and other procedures	
Analyse and solve problems	
Formulate hypotheses from numerical data, graphs, tables	
Use manual skills to set up and operate lab. apparatus	

What procedure do you use to award marks to the students?

Discuss to find out whether marks are awarded by:

impression mark ____________.

or

using a marking scheme ___________.

If by marking scheme, request for a sample exercise and its marking scheme.

How do you determine the criteria for assessment of practical exercises or tests?

Carry out a discussion and find out if any of the following constitute the criterion for award of marks during practical assessment:

Manipulative skills	
Carrying out observational investigation	
Following instructions	
Presenting and interpreting results	
Planning investigations practically	

Space for noting down any relevant issues that might arise in the course of discussion.

3. What are some of the problems biology teachers encounter in implementing the CA Scheme?

Lead question:

Are there any particular problems that you encounter while carrying out Continuous Assessment?

	Problem	Extent of problem	
		Major	Minor
Availability of time			
Preparation of exercise			
Preparation of tests			
Marking exercises/tests			
Entering marks on CA record forms			
Co-operation from Head of School			
Student absenteeism			
Co-operation from Examination Council			
Availability of CA record forms			
Co-operation from students			
Teacher/pupil relationship			
Others:			

Discuss further the nature of the problems and the extent of the problem.

4. What are the biology teachers opinion on the overall conduct of Continuous Assessment?

<u>Lead question</u>:

How would you describe the overall conduct of CA? (For example fairness to pupils, teacher etc.)

____

____

____

____

____

In your opinion should CA continue or should it be stopped?

CA to continue _______________.

CA to be stopped _______________.

If CA is to continue what changes would you like to see made in the present system?

____

____

____

____

Other comments:

5. What is the relationship between the CA score and the final examination score?

Lead question:

What has been the relationship between the CA score you award your students, and the final examination score? (for example, Do those you grade high usually get high scores in the final examination?)

Space for recording response and comments:

Appendix Va: The Summary of the Examiners' classification of the Theory examination questions and items according to Ability Area being tested by the question/item.

	Objective	Question/Item No.
	Ability to	
1.	Show knowledge and understanding of biological facts, concepts, principles and theories.	All items
2.	Comprehend biological texts.	4; 5; 6; 7; 8; 9; 11; 36; 37; 38; 39; 40; 59; 60; 61; 62; 63; 75.
3.	Apply biological principles to real life experience.	1; 2; 4; 5; 6; 34; 35; 62; 63; 72; 73; 74; 75.
4.	Plan experiments.	-
5.	Follow instructions in carrying out experimental and other biological procedures.	-
6.	Make and record observations accurately.	55; 56; 57; 58; 59; 60; 61; 62.
7.	Interpret data to draw conclusions and make inferences.	5; 9; 36; 37; 38; 39; 40; 55; 56; 57; 58; 59; 60; 61; 62; 63.
8.	Draw biological diagrams accurately.	-
9.	Analyse and solve problems.	5; 9; 34; 36; 37; 38; 39; 40.
10.	Formulate hypotheses from numerical data, graphs, tables and other evidences.	34.
11.	Explain and discuss.	1; 2; 3; 34; 35; 36; 37; 38; 39; 40; 41; 42; 43; 44; 45; 46; 47; 48; 49; 50; 51; 52; 53; 54; 55; 56; 57; 58; 59; 60; 61; 62; 63; 64; 65; 66; 67; 68; 69; 70; 71; 72; 73; 74; 75.
12.	Write accurate reports of experiments and other biological procedures.	-
13.	Use manual skills to set up and operate laboratory apparatus.	-

Appendix Vb: The summary of the Examiners' Classification of the Practical examination questions/items according to Ability Area being tested by the question/item.

	Objective	Question/Item No.
	Ability to:	
1.	Show knowledge and understanding of biological facts, concepts, principles and theories.	1; 2; 3; 4; 5; 6; 7; 8; 9; 10; 11; 12; 13; 14; 15; 16; 17; 18; 19. (All items).
2.	Comprehend biological texts.	-
3.	Apply biological principles to real life experiences.	-
4.	Plan experiments.	1.
5.	Follow instructions in carrying out experimental and other biological procedures.	1; 14; 15; 16; 17; 18; 19.
6.	Make and record observations accurately.	1; 12; 15; 16; 17; 18; 19.
7.	Interpret data to draw conclusions and make inferences.	1; 2; 3; 4; 5; 19.
8.	Draw biological diagrams accurately.	12.
9.	Analyse and solve problems.	1.
10.	Formulate hypotheses from numerical data, graphs, tables and other evidences.	-
11.	Explain and discuss.	-
12.	Write accurate reports of experiments and other biological procedures.	1.
13.	Use manual skills to set up and operate laboratory apparatus.	1; 13; 15.

Appendix VIa: The Summary of the Examiners' Classification of the Theory examination questions and items according to Topic Area being tested by the question or item.

	Curriculum Topic	Question/items No.
1.	Classification.	7; 14; 16; 18; 20; 22; 25; 27; 28;57.
2.	Cell structure and function.	8; 11; 22; 27; 28.
3.	Soil and plant growth.	1; 2; 5; 70.
4.	Nutrition in plants and animals.	5; 19; 30; 32; 33; 34; 35; 36; 37; 38; 39; 40; 45; 48; 74.
5.	Respiration in plants and animals.	3; 12; 31; 41; 44; 46; 59; 60.
6.	Transport system in plants and animals.	8; 9; 11; 17; 33; 44; 55; 56; 57; 58; 62; 63; 67; 73; 74.
7.	Excretion and Osmoregulation in mammals.	32; 38; 39; 49; 51; 72.
8.	Control and co-ordination systems in organisms.	6; 13; 15; 42; 43; 49; 52; 72.
9.	Movement.	26; 28.
10.	Growth and development.	14; 23; 27; 28; 54; 66; 75.
11.	Reproduction.	2; 23; 24; 29; 47; 75.
12.	Genetics.	2; 21; 24; 29; 69.
13.	Evolution.	2; 7; 20; 25; 50.
14.	Ecology and interdependence of organisms.	1; 2; 4; 7; 53; 68.
15.	Parasites and diseases of man.	2; 10; 64; 65.
16.	Man and his natural resources.	1; 2.

Appendix VIb: The Summary of the Examiner's classification of the Practical examination questions and items according to Topic Area being tested by the question/item.

	Curriculum Topic	Question/item No.
1.	Classification.	6; 8; 10; 14.
2.	Cell structure and function.	10.
3.	Soil and plant growth.	-
4.	Nutrition in plants and animals.	1; 2; 3; 4; 5; 8; 9; 12; 14; 15; 16; 17; 18; 19.
5.	Respiration in plants and animals.	-
6.	Transport system in plants and animals.	12; 13.
7.	Excretion and Osmoregulation in mammals.	6; 7.
8.	Control and co-ordination systems in organisms.	-
9.	Movement.	-
10.	Growth and Development.	17.
11.	Reproduction.	14; 15; 16; 17; 18; 19.
12.	Genetics.	-
13.	Evolution.	-
14.	Ecology and interdependence of organisms.	9; 11.
15.	Parasites and diseases of man.	-
16.	Man and his natural resources.	-