DOCUMENT RESUME

ED 184 877 SE 030 547

AUTHOR	Hofstein, Avi; Giddings, Geoffrey
TITLE	Trends in the Assessment of Laboratory Performance in High School Science Instruction. Technical Report No. 20.
INSTITUTION	Iowa Univ., Iowa City. Science Education Center.
PUB DATE	Feb 80
NOTE	38p.
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Academic Achievement; Educational Research; Evaluation Methods; *Laboratory Experiments; Science Education; Science Laboratories; *Scientific Principles; Secondary Education; *Secondary School Science; *Student Evaluation

ABSTRACT

Presented is a review of three methods for evaluating the effects of laboratory work in high school science instruction. The report focuses specifically on evaluating student performance, what aspects are to be evaluated, and how to evaluate. A table is given of some of the limitations and benefits of each of the assessment procedures. The authors list several recommendations for use of the different types of assessments. (SA)

****

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

****

ERIC

ED184877

SCIENCE EDUCATION CENTER

The University of Iowa

February 1980

SE 030 547

technical report 20

Trends in the Assessment of Laboratory Performance in High School Science Instruction

by

Avi Hofstein and
Geoffrey Giddings

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Robert Yager
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

The Technical Report Series

The Technical Report Series of the Science Education Center, University of Iowa, was established by action of the faculty during 1973. The series provides a mechanism for communicating results of research, developmental projects, and philosophical investigations to others in Science Education. The reports include details and supporting information not often included in publications in national journals.

Authors of technical reports include the faculty, advanced graduate students, alumni, and friends of science education at Iowa. Technical reports are distributed to all major Science Education Centers in the United States. Reports are also generally available upon request for the cost of packaging and mailing.

Major programs centered in Science Education at the University of Iowa include the following: Science Foundations, a core course in Liberal Arts for undergraduates in education; a special concentration in science for elementary education majors; an undergraduate and a graduate sequence in the history and philosophy of science; a general science major in Liberal Arts, including five emphases for secondary science teaching (biology, chemistry, earth science, environmental studies, and physics); Iowa-UPSTEP, a model six year sequence for preparing new science teachers at the secondary level; undergraduate and graduate programs in environmental studies; Iowa-ASSIST, a statewide curriculum implementation program for in-service teachers; SSTP, a summer and academic year program series for highly interested and motivated secondary school students; self-instruction materials, including computer-based programs.

Major research thrusts at Iowa not reflected in the listing of special programs include: Piagetian Developmental Psychology, Classroom Interaction Studies, Teacher Skills and Attitudinal Studies, Effects of Individual Differences on Learning Science, Philosophical Studies, and Simulation Methods.

Information concerning the Technical Report Series can be received by contacting the Science Education Librarian, Room 470, Science Education Center, Univesity of Iowa, Iowa City, Iowa 52242. Lists of dissertation and thesis reports are available. Also, Field Service Reports, Special Iowa-ASSIST Reports, Special Reports concerning SEC Progress, reports of faculty research, and material describing the various facets of the programs at Iowa are available form the same source.

Since the primary function of the Technical Report Series is communication, comments from you and other consumers of the series are solicited.

Robert E. Yager, Coordinator
Science Education Center
University of Iowa

ERIC

THE TECHNICAL REPORT SERIES

No. 1 - "Newton's Law: A Computer-Based Simulation for Introductory Physics" by Vincent N. Lunetta

No. 2 - "Some Considerations in the Development, Implementation and Evaluation of Curricula" by David Cohen

No. 3 - "Research Related to the Work of Jean Piaget: A Critique" by Darrell G. Phillips

No. 4 - "SCIENCE FOUNDATIONS: A Science Program for the Non-Science Student" by James A. Shymansky

No. 5 - "Computer in the Classroom - A Computer-Based Unit in Teacher Education" by Vincent N. Lunetta

No. 6 - "K-12 Science Education in Iowa" by Lynn W. Glass

No. 7 - "Review of Research on Mathemagenic Behavior: Implications for Teaching and Learning Science" by John T. Wilson and John J. Koran, Jr.

No. 8 - "Baseline Data Concerning Science Teacher Education Programs at the University of Iowa, 1955-73" by Robert E. Yager

No. 9 - "The Iowa-UPSTEP Program in International Perspective" by Pinchas Tamir

No. 10- "The Role of the Laboratory in Science Teaching" by Pinchas Tamir

No. 11- "Mathematical Models of Concept Learning: A Basis for Research in Science Education" by David F. Treagust and Vincent N. Lunetta

No. 12- "A Computer Program Designed to Identify Behavior Patterns in Observational Data" by James A. Shymansky, John E. Penick, and Jay D. Wortman

No. 13- "The Cooperating School in Teacher Education--Source of Theory or Practice?" by Robert J. Hardingham

No. 14- "A Laboratory Structure and Task Analysis Inventory" by Marlene Fuhrman, Vince Lunetta, Shimshon Novick, and Pinchas Tamir

No. 15- "Science Laboratory Interaction Categories (SLIC): An Instrument to Observe and Analyze the Various Parameters of Activity-Centered and Vocational Education Classrooms" by William C. Kyle, Jr., John E. Penick, James A. Shymansky

No. 16- "A Comparative Study of Instructor and Student Behaviors and the Role of the Instructor in Introductory and Advanced Laboratories of Five Science Disciplines at the University Level" by William C. Kyle, Jr., John E. Penick, James A. Shymansky

No. 17- "A Formative, Descriptive Evaluation of the Iowa-UPSTEP Model" by John E. Penick

No. 18- "A Survey of Perceptions of Major Professional Problems and Recommendations for Their Solution by Science Educators Representing Major Centers in the U.S." by Robert E. Yager

No. 19- "The Development and Maintenance of a Unique and Effective In-Service Teacher Education Model" by Daniel S. Sheldon

Trends in the Assessment of Laboratory Performance

in High School Science Instruction

Avi Hofstein
Visiting Professor
from
Department of Science Teaching
The Weizmann Institute of Science
Rehovot, Israel

and

Geoffrey Giddings
Visiting Professor
from
Chemistry Department
Melbourne State College
Carlton, Victoria, Australia

Technical Report 20

Trends in the Assessment of Laboratory Performance

in High School Science Instruction

Table of Contents

INTRODUCTION: THE LABORATORY AND SCIENCE CURRICULA

The history of laboratory work as an integral part of school science learning began early in the twentieth century. The laboratory in the science classroom has long been used to involve students in concrete experiences with objects and concepts. "The laboratory experiment or exercise refers to an instructional procedure in which cause and effect, nature or property of any object or phenomenon is determined by individual experience generally under controlled conditions" (Pella, 1961).

The "new" science curricula in the 60's and 70's resulted in changes in the role assigned to laboratory work. Shulman and Tamir (1973) stated that "the laboratory has always been the most distinctive feature of science instruction. With the advance of new curricula which stress the process of science and emphasize the development of higher cognitive skills, the laboratory has acquired a central role, not just as a place for demonstration and confirmation but rather as the core of the science learning process."

Similar ideas are shared also by other science educators like Schwab (1962), Hurd (1969), Hincksman (1974), and Tamir and Lunetta (1978). They have expressed the view that the major uniqueness of the laboratory lies in providing students with opportunities to engage in processes of investigation and inquiry. As a result, in 1970 the NSTA Commission of Professional Standards and Practices thought the case for school science laboratories too obvious to need much argument. It was stated that:

> "The time is surely passed when science teachers must plead the case for school laboratories. It is now widely recognized that science is a process and an activity as much as

it is an organized body of knowledge and that, therefore, it cannot be learned in any deep and meaningful way by reading and discussion alone."

EVALUATING STUDENT PERFORMANCE

Since effective testing and evaluation is considered to be one of the most needed areas of improvement in science teaching (Lee, 1969; and Golman, 1975), there is a need to develop creative methods to evaluate laboratory instruction. Grobman (1970) observed that with few exceptions, evaluation (of students) has depended on written testing:

> "There has been little testing which requires actual performance in a real situation, or in a simulated situation which approaches reality . . . to determine not whether the students can verbalize a correct response but whether he can perform an operation, e.g. a laboratory experiment or an analysis of a complex problem. . . . This is an area where testing is difficult and expensive, yet since in the long run, primary aims of science projects generally involve doing something rather than writing about something, this is an area which should not be neglected."

These views find support in research evidence. Ben-Zvi, et al. (1977) found that teachers' assessment of their students in the sciences is mainly based on students' performance in cognitive (pencil-and-paper) tests. Robinson (1969) in the United States found that a low correlation (0.33) exists between laboratory practical examinations and written paper-and-pencil tests.

This concern was also raised by science educators in the United Kingdom. According to Kelly and Lister (1969):

> "Practical work involves abilities, both manual and intellectual, which are in some measure distinct from those used in non-practical work."

These arguments and other findings lead one to the proposition that there is a need to develop special assessment inventories to evaluate student performance in the science laboratory.

WHAT CAN WE EVALUATE?

In order to assess the various outcomes of laboratory work there is a need to identify the component skills associated with laboratory work. In the United States, Jeffrey (1967) suggested six areas associated with laboratory work:

1. Communication: identification of laboratory equipment and operations;
2. Observation: recording of observations and detecting errors in techniques:
3. Investigation: accurate recording of measurable properties of an unknown substance;
4. Reporting: maintenance of a suitable laboratory record;
5. Manipulation: skills in working with laboratory equipment;
6. Discipline: maintenance of an orderly laboratory and observation of safety procedures.

Jeffrey stressed the need to design practical examinations in which the students will be involved in manipulating materials and apparatus.

Kampa and Ward (1975) in the United Kingdom described the overall process of practical work in science education as involving four phases:

1. Planning and designing investigation in which the students predict results, formulate hypotheses, and design procedures;
2. Carrying out experiments in which the student makes decisions about investigative equipment;
3. Observations of a particular phenomenon.

4. Analysis, application and explanation in which the student is processing data, discussing results, exploring relationships and formulating new questions.

The Joint Matriculation Board (JMB, 1979) and the University of London Examination Board (1977) concentrated their assessment of practical work on five main areas:

TABLE 1

Assessment Areas	Relative Weighings
1. Manipulative skills	25% - 30%
2. Skill in observation and the accurate recording of such	25% - 30%
3. Ability to interpret observations	20% - 25%
4. Ability to devise and plan experiments	10% - 15%
5. Attitudes	10% - 20%

HOW CAN WE EVALUATE?

Three methods of evaluation available to the science teacher for the assessment of laboratory work are:

1. Written evidence - either traditional laboratory reports or paper-and-pencil tests;

2. One or more practical examinations;

3. Continuous assessments by the science teacher.

The form(s) of assessment utilized by the science teacher should take into account the following criteria:

1. Reliability - seen as the degree to which the form(s) of assessment yields consistent and reproducible results;

2. Validity - seen as the degree to which the assessment procedure measures the aims of the laboratory course;

3. Useability - seen as the criterion which must follow once the procedures are reliable and valid for a given tas. This means chosing those which offer the most convenience, economy and interpretability.

A. WRITTEN EVIDENCE

a. Traditionally, science teachers have assessed their students' performance in the laboratory on the basis of written student reports; written during or subsequent to, the laboratory experience. Assessment based solely on the written reports limit the assessable skills to only one of the previously identified skill components, e.g. Jeffrey's (1977) reporting component; Kempa and Ward's (1975) analysis, application and explanation phase and the J.M.B.'s (1979) interpret observation category.

b. The second form of written evidence is the paper-and-pencil test, designed to assess student knowledge and understanding of the use of experimental techniques and of the principles underlying laboratory work and procedures. (See examples in Appendix 1). Kruglak, (1958) suggested two hypotheses for using laboratory-type paper and pencil tests:

1. "It is impossible to measure the creative aspect of laboratory achievement by means of multiple choice type tests." By creative aspect, he means "to see and formulate an experimental problem and develop methods and techniques for solving experimental situations."

2. "It is impossible to measure certain neuro-muscular laboratory skills by means of a paper and pencil test. A student might get a perfect score on a written test but not be able to handle apparatus in the laboratory."

ERIC

Thus it seems that a paper-and-pencil test can assess only two of the above mentioned components (e.g. Jeffrey's (1967) communication and reporting and Kempa and Ward's (1975) planning/design and analysis/application phases). It seems that a paper-and pencil test is capable of measuring the more theoretical component of a laboratory problem. Laboratory items should be included in general science achievement tests but not instead of assessing the student involved with an experiment in a real situation

B. PRACTICAL EXAMINATIONS

Science teachers and those concerned with the examination of laboratory work, have tended to ignore the practical examinations as a means of collecting information on student performance due to problems of implementation and validity.

Ben-Zvi, et al. (1976) conducted a comprehensive study in Israel on the assessment of laboratory outcomes based on the remaining two phases of Kempa and Ward's (1975) outline, namely the phase concerned with the conducting of, and decision making within, the experiment and the observation phase.

Two practical tests were used:

a. Problem solving test: In this test students were asked to design apparatus and experimental method to investigate the following reaction:

$$CdCO_{3(s)} \longrightarrow CdO_{(s)} + CO_{2(g)}$$

The directions for this experiment were given in such a way that students were faced with an open-ended, problem-solving situation. Both planning and actual performance were tested.

b. Manipulative skills test: The method used was that of continuous variation of the following reaction:

$$Pb^{2+}_{(aq)} + CrO_4{}^{2-}_{(aq)} \longrightarrow PbCrO_{4(s)}$$

The students were required to perform experimental work according to well-defined instructions and so the chief purpose of this exercise was to examine manipulative skills.

The individual performance criteria laid down in a checklist embraced four subcategories of manipulative skills suggested by Eglen and Kempa (1974) listed in Table 2:

TABLE 2

A Breakdown Analysis of Manipulative Skills

Skill Components	Generalized Assessment Criteria/Performance Features
1. Experimental Technique	1. Correct handling of apparatus and chemicals; safe execution of an experimental procedure; taking of adequate precautions to insure reliable observations and results.
2. Procedure	2. Correct sequencing of tasks forming part of an overall operation; effective and purposeful utilization of equipment; efficient use of working procedure on the basis of limited instructions.
3. Manual Dexterity	3. Swift and confident manner of execution of practical tasks; successful completion of operation or its consituent parts.
4. Orderliness	4. Tidiness of the working area; good utilization of available bench space; purposeful placing of apparatus and equipment.

Observers were asked to observe students conducting the two practical tests on the basis of a checklist. (See Table 3.)

TABLE 3			
Criteria and Items in the Check List			
Criteria		Yes	No
Technique	Did the student tighten the syringe in the clamp?		
Technique	Is the clamp assembled properly?		
Technique	Did the student connect the glass pipe to the syringe with a rotatory fashion?		
Procedure	Did the student check the sealing of the apparatus after fitting the test tube?		
Manual Dexterity	Was the apparatus assembled efficiently and safely?		
Manual Dexterity	Was there any evidence to show that the student understood what he was meant to do?		
Manual Dexterity	Was the assembly of the apparatus carried out without mishaps?		

Eglen and Kempa (1975) incorporated, in their research study, three different assessment schedules to assess a laboratory sequence.

1. The check list schedule: this is similar to the one used by Ben-Zvi, et al. (1976). See Table 1.

2. The open-ended schedule: this asked the teachers to assess the students subjectively on a 1-5 scale without reference to achievement criteria other than to those listed in Table 1. The following is an example of an open-ended schedule:

> Grade awarded __________
> Major manipulative abilities looked for in the operation:__________
> Features in the student's performance used for the grading:__________

3. The third method is the intermediate schedule. This is similar to the open-ended schedule, but requires teachers to provide separate

assessments for the four subcategories of the manipulation skills domain listed in Table 1. The following is an example of the Intermediate Mode.

1. Methodical working: Grade____
e.g. in the washing of glassware; in the transfer of solid and of solution; in the processes of rinsing and diluting; in the stirring and dissolving operations.
Features in the student's performance used for the grading______________________

____

2. Experimental technique: Grade____
e.g. that used in the handling of glassware and of materials; in the operations of dissolving, transfer, washing and dilution.
Features in the student's performance used for the grading______________________

____

3. Manual dexterity: Grade____
e.g. in the general manner of performance and in the completion of the task.
Features in the student's performance used for the grading______________________

____

4. Orderliness: Grade____
e.g. in the utilization of bench-space and the organization of equipment.
Features in the student's performance used for the grading______________________

____

Kempa and Eglen compared these three methods in order to find out how concordant teachers' assessments of students' manipulative abilities are. They found that 16 open-ended type assessment procedures which are largely based on impression grading led to considerable differences in grades awarded by the teachers for 16 similar practical performances. When assessment is based on specific criteria (like the checklist method), less divergence in the results was observed. But even in this case there was no 100% agreement between the teachers who were involved in the evaluation exercise. It seems that since the "check-list" method is impractical for classroom use, and because it is essentially subjective (relative to the intermediate modes), leads to the results similar to the check list. In the future this method should be utilized by science teachers.

c. Test of obervational ability: Observing scientific phenomena is an important part of the overall process of scientific inquiry conducted in the laboratory. Nevertheless, the direct teaching and assessment of these skills are often neglected in science courses. An observational test was used by Ben-Zvi, et al., (1976). In this test students were asked to perform several simple test tube experiments covering the following perceptual areas: color changes, evolution of gases, temperature changes, and precipitation of solids. Students are asked to record their observations on the following type of sheet:

Experiment I
Add to a test tube 5 ml of solution A and a few grains of Zinc.
Observation 1 . . . .
Observation 2 . . . .
Observation 3 . . . .
Observation 4 . . . .
Observation 5 . . . .

Experiment II
To a test tube add 5 ml solution A and 5 cm long Magnesium Ribbon.
Observation 1 . . . .
Observation 2 . . . .
Observation 3 . . . .
Observation 4 . . . .
Observation 5 . . . .

Another example of a practical examination is by Tamir and Glassman (1971) who developed a laboratory and inquiry-oriented laboratory examination for Israeli biology students. The following were the criteria for this examination (from Tamir, 1974):

1. They should pose some real and intrinsically valuable problems before the students.
2. It should be possible to perform the task and conclude

the investigation within a reasonable time limit, (i.e. 2 hours).

3. The problems should be novel to the examinees, but the level of difficulty and the required skills should be compatible with the objectives of the experiences provided by the curriculum.

4. Since it would be impractical to administer the test individually, it should be administered in a group setting. Since every student will be able to perform only one full investigation, several different problems must be used simultaneously to insure independent work within this group setting. However, for the sake of comparability, the different tests should be convergent on a number of skills with specific weights given to the various responses. It will also be necessary to control for differences in the levels of difficulty as well as for the heterogeneity of variance by employing appropriate statistical moderation procedures.

5. The student performing a complete investigation may encounter certain difficulties at various steps of his work. It is inconceivable that he should fail the whole examination just because, for instance, he made some incorrect observations. Therefore, a procedure is needed for prompting or providing certain leads during the examination without damaging the standards of assessment.

6. Since the tests are based on open-ended problems and measures of divergence are needed, acceptable limits of this divergence must still be set.

7. When tests of this kind are used as external examinations, special logistic problems are involved. For example, while certain materials and laboratory equipment can be prepared by the schools, some materials and organisms must be brought by the examiners in order to prevent the examinees from obtaining clues regarding the tasks to be assigned during the examination. Also, since novelty is an important feature, new problems must be designed every year.

The nature of the examination is illustrated in an example of a problem given to a student (See Appendix 2).

Golman (1975) used both paper-and-pencil and practical test in order to measure the outcomes of laboratory work in an introductory zoology course in college biology. (See Appendix 2 for details.)

C. CONTINUOUS ASSESSMENT

The limitation of practical examinations to those experiments that can be readily administered to students in a limited time obviously restricts both the scope and validity of the assessment. It can also have undesirable effects on the choice of experiments throughout the year; teachers limiting their choice of experiment to those highly related to the type of experiment utilized in a practical test.

More recently there has been a change towards continuous internal assessment of practical skills and abilities conducted by teachers in many school systems. This has been formalized to a large degree in the United Kingdom (JMB, 1979, University of London, 1977) and in Israel (Cohen, 1977, and Cohen, et al., 1978).

This change is a response to a range of concerns expressed by classroom teachers and those responsible for external practical exams. Certainly most teachers would see this task as a natural part of their role as teachers and simply an extension of the assessments that all teachers carry out in their day-to-day teaching.

Continuous assessment on several occasions throughout the year is necessary to cover adequately the variety of tasks and techniques which comprise a total program of practical work.

With a greater involvement in the continuous assessment of practical skills the teacher is likely to develop a greater awareness of the scope and objectives of laboratory work by identifying student strengths that otherwise may not have been reflected in more conventional assessments. Some basic principles of continuous assessment include:

1. Teachers should inform their students at the beginning of the course that their practical work is being assessed in a continuous manner over the whole program. Details

ERIC

should be given regarding the abilities and skills that are to be assessed. Fears have been expressed that the position of the teacher as an assessor may effect adversely the close relationship between student and teacher. Experience suggests that no fear need exist. Teachers are normally involved in the assessment aspect of a student's work as part of their everyday teaching.

2. As far as possible, the teacher should make assessments during a normal practical class and make their assessment procedure as unobstrusive as possible. It is not necessary to assess all students on the same day or on the same experiment. It may be that some students are absent or more likely there will be insufficient time to assess all students. Also any one experiment is unlikely to provide the opportunity to assess all the areas at the one time.

3. There are three main ways in which marks can be allocated to a particular objective and teachers will probably find that they will have to use all of them at some stage of the program:

   i) A Mark Scheme. This will be most useful when marking written evidence of observation, interpretation, planning and accuracy.

   ii) Marking by Impression on a Single Occasion. This will be useful for marking evidence that is less precise than in (i). For example, a teacher may wish to assess dexterity in handling unfamiliar apparatus - say in some simple paper chromatography. Teachers should try to assess only one such quality during the session by impression using some of the rating scales suggested (See Appendix 3).

   iii) Marking by Impression over a Period of Time. This will apply mainly to attitudes to practical work, but some of the less precise aspects of manipulative skills may also be better assessed periodically rather than in single experiments. The period can be once a semester, once a year, or, for some objectives, once at the end of the course.

The students are assessed on the following five main areas:

(See Table next page.)

TABLE 1	
Assessment Areas	Relative Weighings
1. Manipulative skills	25% - 30%
2. Skill in observation and the accurate recording of such	25% - 30%
3. Ability to interpret observations	20% - 25%
4. Ability to devise and plan experiments	10% - 15%
5. Attitudes	10% - 20%

Scales for such assessment are given in detail in Appendix 3. Appendix 4 contains two examples of record sheets for the tabulation of such measures.

SUMMARY

In this report three methods for assessing outcomes of laboratory work in the context of high school science instruction are reviewed. Two of the methods, practical examination and continuous assessment, assess a student's performance while he is conducting an experiment, the other relies on two forms of written evidence--traditional report writing and paper-and-pencil tests.

Since assessment based solely on written evidence assesses only a limited range of laboratory skills, this type of assessment procedure should be used only as a part of the total assessment.

The second method reviewed is the practical examination in which students are assessed on the basis of certain previously identified criteria "criterion referenced testing", (Popham and Husek, 1969).

The third method reviewed is a continuous laboratory-based student assessment in which the student is assessed by his own teacher over a number of different classroom experiments during a particular course.

The following comparison table reflects some of the limitations as well as the benefits of each of the assessment procedures:

TABLE 5

COMPARISON OF THE THREE METHODS OF LABORATORY ASSESSMENT IN THE CONTEXT OF HIGH SCHOOL SCIENCE INSTRUCTION

DIMENSION	WRITTEN REPORT	PAPER-AND-PENCIL TEST	PRACTICAL EXAMINATION	CONTINUOUS ASSESSMENT
Outcomes Measured	Comprehension & Interpretation	Comprehension Interpretation & Planning	Manipulation & Comprehension	Can Cover All the Objectives in the Practical Domain
Reliability	Very Low	Low-Medium	Very Low	High
Validity	Low	Low	Low-Medium	Medium-High
USABILITY:				
Convenience	High	High	Low	Medium
Economy	High	High	Low	Low
Interpretability	Low	Low	Low	High
Assessor	Teacher	Teacher	Teacher or External Assessor	Teacher
Degree of Involvement of Assessor	Low	Low	High	High
Level of Anxiety on Behalf of Student	Low-Medium	Medium	Very High	Low-Medium
Use in Redesigning & Planning of Future Program	Low	Low	Low	High
Number of Topics Covered	Limited	Large Number	Limited	Large Number
Feedback to the Teacher	Occasional	Continuous	Occasional	Continuous

Apart from the highly subjective nature of the continuous assessment method, this procedure of assessment is one teachers should be encouraged to examine more closely. Since it was found (Cohen, 1977, and Cohen, et al., 1978) that student's ability to manipulate laboratory equipment in a given science experiment is directly related to his understanding of the principle underlying the experiment, teachers might consider using both paper-and-pencil and continuous assessment to evaluate student's aoilities in the science laboratory.

It may be concluded that, if one agrees that practical work in the laboratory has a unique role in science instruction, then the assessment of laboratory work should not be overlooked by the teacher or by external examining bodies.

ERIC

REFERENCES

Ben-Zvi, R., Hofstein, A., Samuel, D., and Kempa, R.F. The effective ness of filmed experiments in high school chemical education, Journal of Chemical Education, 53, 518-520, 1976.

Ben-Zvi, R., Hofstein, A., Samuel, D., and Kempa, R.F. Modes of instruction in high school chemistry. Journal of Research in Science Teaching, 14, 433-439, 1977.

Cohen, I. The analysis and assessment of laboratory work in the "chemistry for high school" curriculum. Unpublished M.S. Thesis. The Weizman Institute of Science, Rehovot, Israel. 1977 (In Hebrew).

Cohen, I., Ben-Zvi, R., Hofstein, A., and Samuel, D. The assessment of practical work in high school chemistry instruction, Halaha-L'Masse. Theory and Practice in Curriculum Instruction, 2, 131-143, 1978.

Eglen, J.R. and Kempa, R.F. Assessing manipulative skills in practical chemistry. School Science Review, 56, 261-273, 1974.

Golman, M. Assessing laboratoy instruction in biology. Iowa Science Teachers Journal, 12, 4-10, 1975.

Grobman, H. Developmental Curriculum Projects: Decision Points and Processes, Itasca, Illinois, Peacock Publishers, Inc., 1970.

Hincksman, N.O., The function of the school laboratory, The Australian Science Teachers Journal, 19, 81-86, 1973.

Hofstein, A. The Measurement of the effectiveness of laboratory work in high school chemistry instruction in Israel. Unpublished Ph.D. Thesis, The Weizman Institute of Science, Rehovot, Israel. 1975 (in Hebrew).

Hurd, P.D. New Directions in Teaching Secondary School Science. Chicago, Rand McNally and Co. 1969.

Jeffrey, J. Identification of objectives of the chemistry laboratory and development of means for measuring student achievement of some of these objectives. Unpublished doctoral thesis, The University of Texas, (University Microfilms), Ann Arbor, Michigan, No. 66-1928, 1967.

J.M.B., (Joint Matriculation Board) The Internal assessment of practical skills in chemistry (advanced) - suggestions for practical work and advice on sources of information. Universities of Manchester, Liverpool, Leeds, Sheffield and Birmingham, January 1979.

Kelly, P.J. and Lister, R.E. Assessing practical ability in Nuffield A-level biology in J.F.E. Eggleston and J.F. Kerr (Eds.), Studies in Assessment, London, English Universities Press, 1969.

Kempa, R.F. and Ward, J.F. The effect of different modes of task orientation on observations attained in practical chemistry, Journal of Research in Science Teaching, 12, 69-76, 1975.

Kruglak, H., Evaluating laboratory instruction by use of objective type tests. American Journal of Physics, 26, 31-32, 1958.

Lunetta, V.N. and Tamir, P., An analysis of laboratory activities in two modern science curricula: Project Physics and PSSC. National Association for Research in Science Teaching, Toronto, Ontario, 1978.

NSTA Commission on Professional Standards and Practices: "Conditions for Good Science Teaching in the Secondary School", National Science Teachers Association, Washington, D.C., 1970.

Pella, M.D. The laboratory and science teaching, The Science Teacher, 28, 29-31, 1961.

Popham, J.W. and Husek, T.R. Criterion referenced testing, Journal of Educational Measurement, 6, 1-3, 1969.

Robinson, J. Evaluating laboratory work in high school biology, American Biology Teacher, 31, 236-240, 1969.

Schwab, J.J. The teaching of science as inquiry, In J.J. Schwab and B.F. Brandewein, The Teaching of Science, Cambridge, Massachusetts, Harvard University Press, 1962.

Shulman, L. and Tamir, P. Research on Teaching in the Natural Science in Second Handbook of Research on Teaching, R. Travers, (Ed.), Rand-McNally, and Co., Chicago, Ill., 1098-1148, 1973.

Tamir, P. and Glassman, F. A practical examination for BSCS students, Journal of Research in Science Teaching, 7, 107-112, 1970.

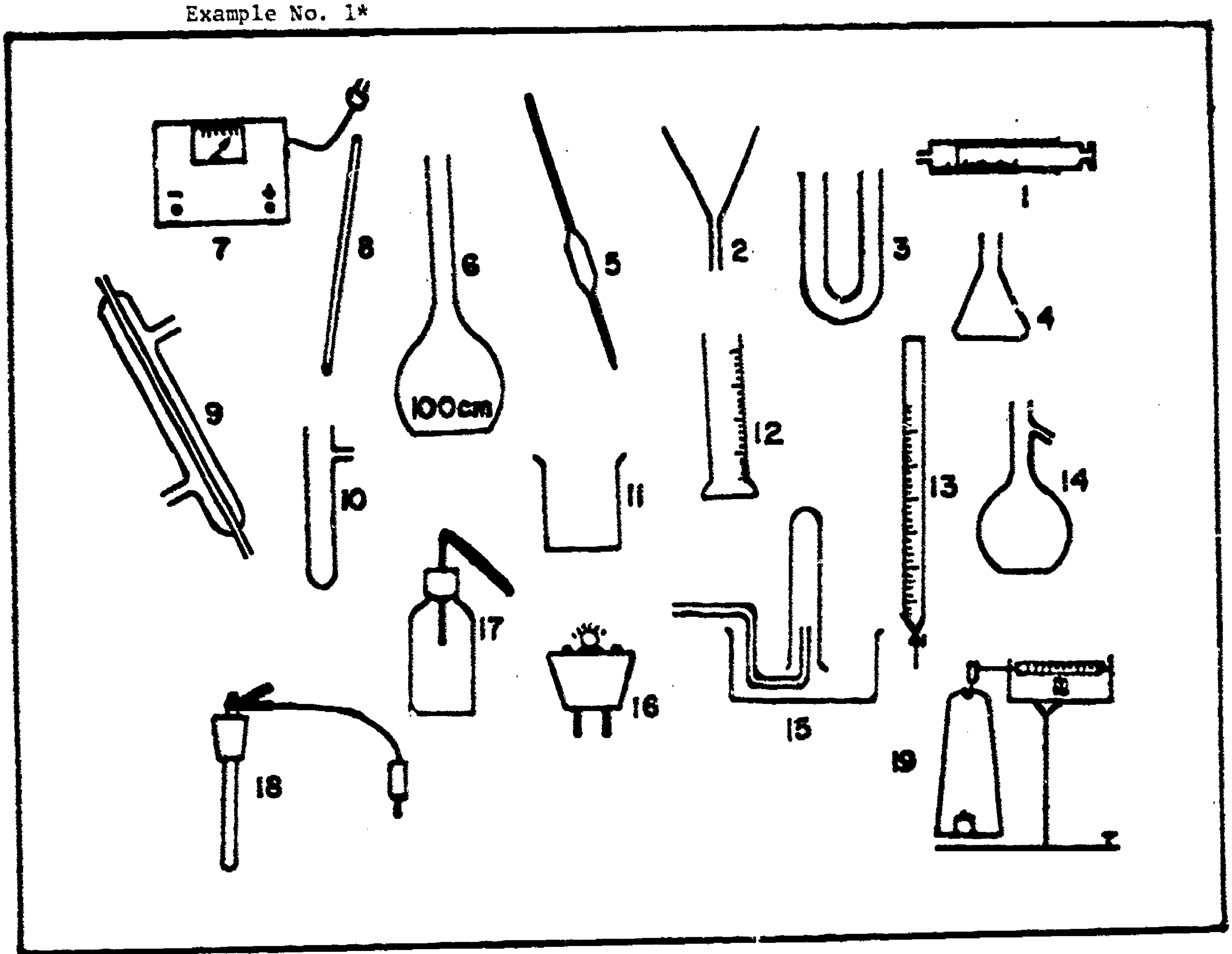
Tamir, P. An inquiry oriented laboratory examination, Journal of Educational Measurement, 11, 25-33, 1974.

University of London, (G.C.E.), Notes of Guidance and Report Form for the Optional Internal Assessment of Special Studies, University of London, 1977.

Appendix 1

Examples of Paper-and-Pencil Tests

Example No. 1*



Which of the following will be used for:

A. Distillation . . . .

B. Measurement of electrical conductivity . . . . . .

C. Preparation of 0.1M NaOH . . . .

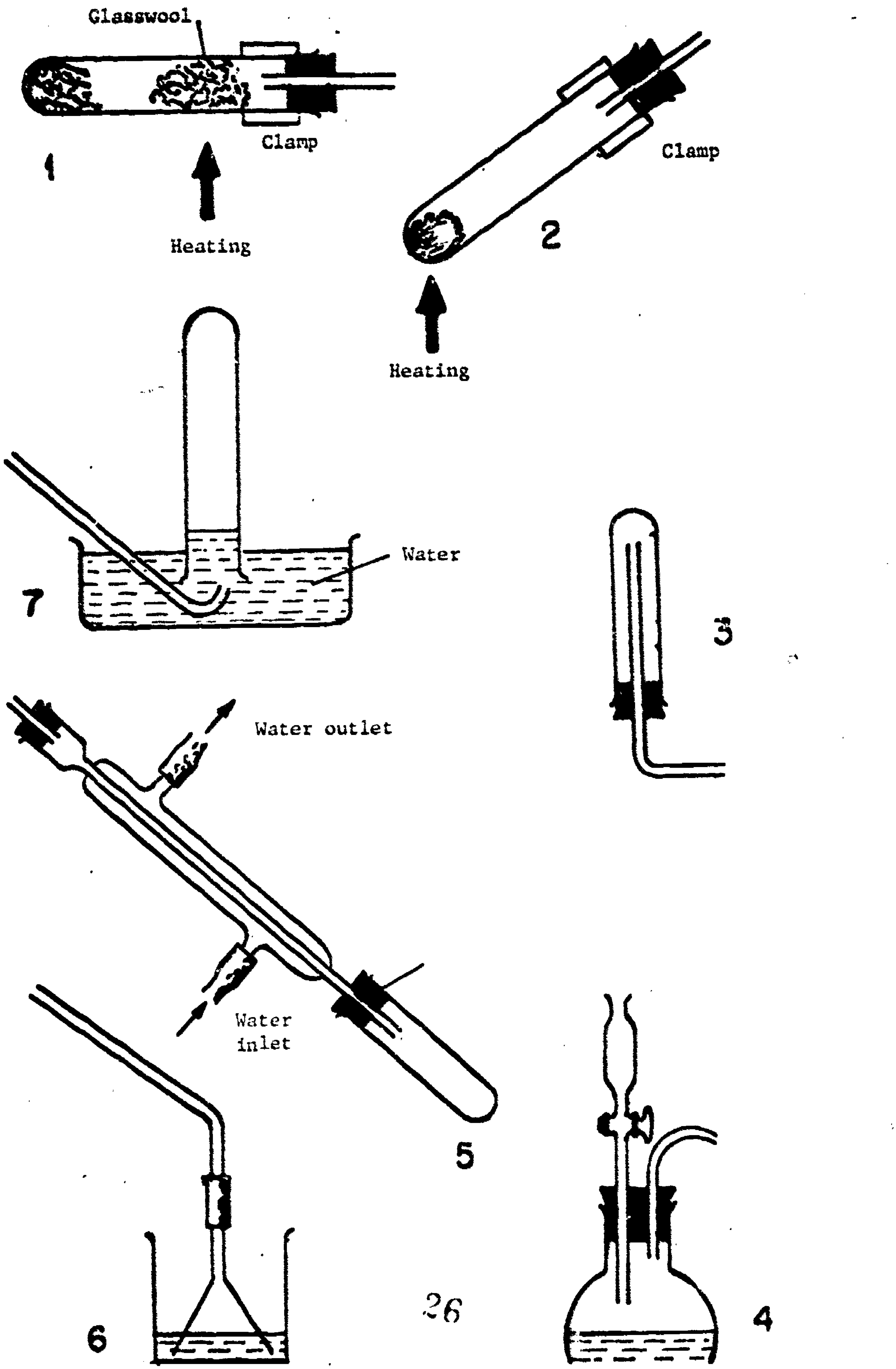
D. Preparation of H_2 . . . .

E. Preparation of Cu from $Cu^{2+}_{(aq)}$ . . . .

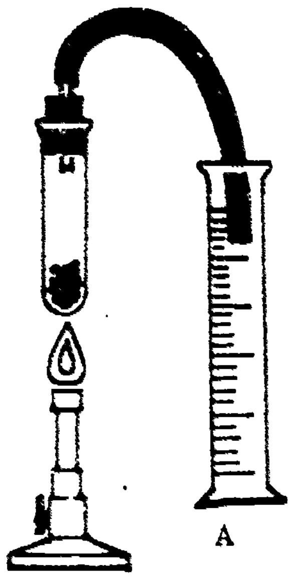
*From: Hofstein (1975)

ERIC

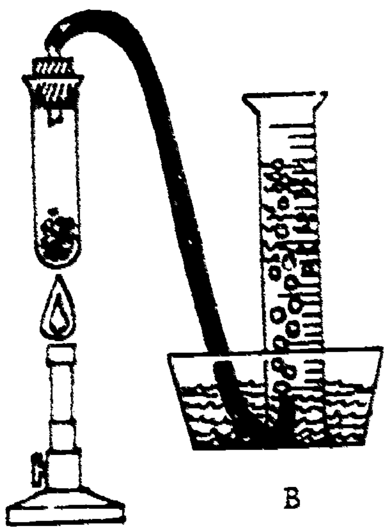
Example No. 2: Which of the below is the best arrangement to collect a gas which does not dissolve in water?



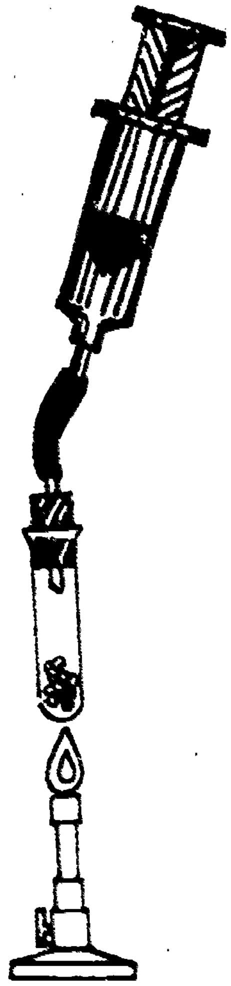
Example No. 3:* A solid MCO_3 was heated and gas evolved. Which of the following arrangements is the best method to measure the volume of the gas evolved?



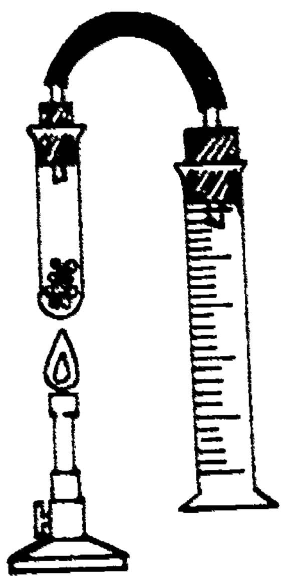
A



B



C

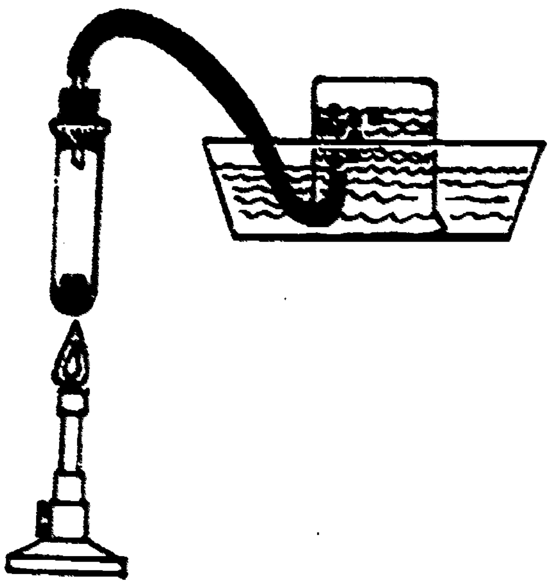


D

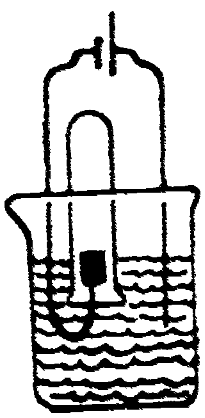
* From Cohen, (1977).

Example No. 4:* In the following questions indicate whether the arrangement made is appropriate or not.

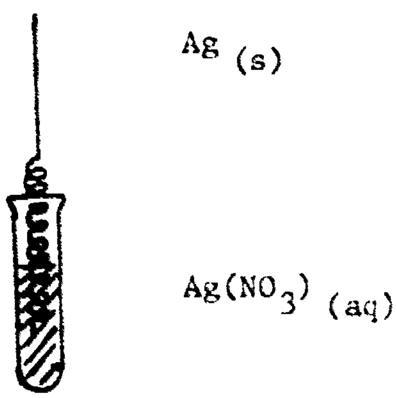
1. Is the following arrangement appropriate to collect NH_3 from $(NH_4)_2\ (HPO_4)_2\ _{(s)}$?



2. Preparation of $H_{2\ (g)}$ from $H_2SO_{4\ (aq)}$.

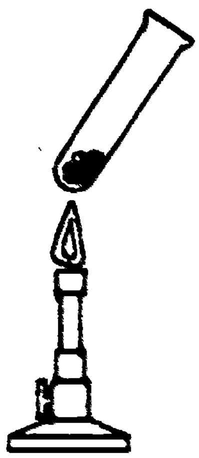


3. Preparation of Ag(s) from $AgNO_{3(aq)}$.



* From Cohen, (1977).

4. For $HgO_{(g)}$ $\xrightarrow{\text{Heat}}$ $Hg_{(1)} + 1/2\ O_{2\ (g)}$.



5. For NaCl $\xrightarrow{\text{Heat}}$ $Na_{(s)}$ + $1/2\ Cl_{2\ (g)}$.

Appendix 2

Example I: Practical Test in Biology*

Problem 1.

Materials: yeast, buffered carbonate solution, 0.2% neutral red unlabelled; NaOH, NH_4OH , (all 0.1 M); (test tubes, flasks, etc.).

When the student enters the laboratory, he finds the materials and the examination paper. He is requested at several points during the examination to call the examiner who will circle specific answers with red ink and hand over cue notes that direct the continuation of the investigation. These notes are used in order to avoid or reduce the discussion time between the examiner and the examinee.

Examination Paper:

1. On the table you will find a red liquid. Using the materials on your table, find out the color of this liquid in various pH concentrations. Write down the procedures and the results.
2. Prepare a yeast suspension (detailed instructions are given).
3. Mix 25 ml of the yeast suspension with an equal amount of the red liquid. Observe for five minutes. Record your observations.
4. Design a control to the test performed in No. 3 and show it to the examiner.Observe for five minutes. Record you observations.
5. Perform a controlled test and record the results.
6. Explain your observations.
7. Suggest a way to test whether your explanation is correct. Show your design to the examiner. (Examiner circles the answer and hands over a cue-note with instructions.)
8. Perform the test and write down your results and conclusions.
9. Do the results support your suggested explanation? If not, suggest another explanation.

* From Tamir, (1974).

10. In any case, suggest an additional way to test your explanation.

11. Carry out the test and record the results and conclusions.

12. Do you think that the yeast you worked with were alive? Give reasons.

13. Design an experiment to test your answer to No. 12. (Examiner will hand over a cue-note.)

14. Perform the experiment and record the results.

15. Briefly sum up the findings and conclusions of your investigation

The examiner gets a sheet with detailed instructions regarding his behavior during the examination as well as sample answers. For instance, sample answer 15 reads:

> Yeast cell membrane permits absorption and accumulation of neutral red. Since the pH in the yeast is acid, the color of the cell is red. The buffer solution is basic and therefore has a yellow color. When the yeast cells are dead, they lose their selective permeability, their content mixes with the buffer solution and hence the yellow color dominates.

Marking is done according to a predetermined key. The relative weight assigned to each skill which is the same for each of the test problems is as follows (in percents): manipulation --10, self-reliance--10, observation--15, experimental design--25, reporting and communication--15, reasoning--25.

In this problem, presented above, manipulation is measured by items 1, 2, 3, 5, 8, 11, and 14; observation by items 1, 3, 5, 8, 11, and 14; self reliance and experimental design by items 4, 7, 10, 13; communication by items 5, 8, 11, and 14; and reasoning by items 6, 8, 9, 11, 12, and 15.

Example II: Laboratory Pretest Items*

1. Objective: Given a Celsuis thermometer and a flask containing a liquid, the student will be able to determine the temperature of the liquid to the nearest degree.

   Laboratory Station Set-Up: 500 ml flask containing water and a Celsius thermometer.

   Question: What is the temperature of this solution in degree Celsius?

2. Objective: Given an object in balance on the triple beam balance, the student will be able to determine the weight of the object to the nearest 0.1 g.

   Laboratory Station Set-Up: 100 ml beaker on the triple beam balance and balanced.

   Question; What is the weight of this beaker in grams?

3. Objective: Given a graduated cylinder containing a liquid, the student will be able to determine the volume of the solution to the nearest ml.

   Laboratory Station Set-Up: 100 ml graduated cylinder containing a light blue colored water solution up to the 53 ml level.

   Question: How many milliliters of solution is contained in this graduated cylinder?

4. Objective: Given an object and a millimeter ruler, the student will be able to determine the length of the object to the nearest millimeter.

   Laboratory Set-Up: Plain microscope slide and millimeter ruler.

   Question: What is the length of this microscope slide in millimeters?

5. Objective: Given several fruit flies with different eye characteristics, the student will be able to sort these flies into groups with identical eye characteristics.

   Laboratory Station Set-Up: Four freshly killed fruit flies in focus on the states of a stero-microscope with a fly having sepia eyes and other traits wild type labeled "A", a fly having wild type eyes and other traits wild type labeled "B", a fly having white eyes and other traits wild type labeled "C", and a fly having sepia eyes and other traits wild type labeled "D".

   Question: Which two flies are most similar in appearance if size is disregarded?

*From Golman,(1975).

6. Objective: Given several pieces of glassware for measuring, the student will be able to select the piece of glassware which will provide the most accurate measurement of a solution.

   Laboratory Station Set-Up: 10 ml volumetric pipette labeled "A", 10 ml graduated cylinder labeled "B:, 10 ml graduate pipette labeled "C", and 10 ml graduated beaker labeled "D".

   Question: Which piece of laboratory glassware would be best for measuring 9.5 ml of a solution most accurately?

7. Objective: Given several labeled flasks as part of an experimental design, the student will be able to select the control and experimental flasks.

   Laboratory Station Set-Up: Four flasks containing the following solutions and labels: A - yeast cells, water and glucose, B - yeast cells, water and sucrose, C - yeast cells, water and starch, D - yeast cells and water.

   Question: If flask "A" is an experimental flask, which flask is the control in this experiment?

8. Objective: Given a microscope with the 10X ocular and 5X objective in position, the student will be able to determine the magnifying power of this lens system.

   Laboratory Station Set-Up: Microscope with 10X ocular and 5X objective in position with slide on microscope stage.

   Question: If you were to examine a specimen on a slide through this microscope, how many times larger would the specimen appear with this lens system in position?

9. Objective: Given the diameter of a microscopic field in millimeters, the student will be able to determine the diameter of the field in microns.

   Laboratory Station Set-Up: None.

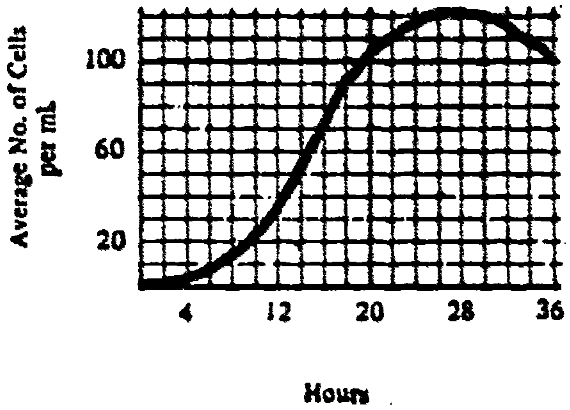
   Question: The diameter of a microscopic field was measured and found to be 1.2 millimeters. What is the diameter of the field in microns?

10. Objective: Given a graph illustrating cell population growth, the student will be able to determine the period of most rapid population growth to the nearest hours.

    Laboratory Station Set-Up: None.

Question: Examine the graph below illustrating cell population growth. The period of most rapid population growth occurs between the hours of -

A. 4 and 12
B. 16 and 24
C. 24 and 32
D. 32 and 36



Appendix 3

Examples of Criteria and Rating Scales for Continuous Assessment*

The following criteria are suggested for marking on a ten-point scale. For each of the abilities the criteria have been specified in terms of ranges of two marks; this procedure is suggested because to attempt to distinguish between single marks would results in criteria too complex for teachers to handle effectively.

Example A. The possession of appropriate manipulative skills.

10-9** Good all-round ability to carry out full range of skills. Intuitively does the right things. Good appreciation of precision of apparatus. Quantitative results within the expected range.

8-7 Good general ability but limited in certain skills.

6-5 Routine worker who does not vary the degree of precision in accordance with the requirements of the particular situation. Does not always appreciate delicacy of apparatus.

4-3 Rather careless in handling of apparatus. Sometimes fails to follow experimental instructions.

2-1 Careless in handling apparatus. Quantitative results not acceptable.

Manipulative skills can be assessed by marking the results of experiments, for example, yield, purity and accuracy (accuracy of work rather than accuracy of calculation); they can also be assessed by direct observation of such things as orderliness, methodical working, dexterity and speed. It is desirable that this assessment should not be confined to routine operations; it should include the ability of students to adapt their manipulative skills to new situations and to follow unfamiliar instructions.

---

* Based on documents produced by the JMB (1979) and The University of London Examination Board (1977).

** Suggested rating

It may be possible for teachers to use specific experiments for assessing some parts of manipulative skill, particularly yield, purity and accuracy. However, many teachers feel that a grade based on a general impression of work over a period of say a semester or even a year is more appropriate for the less exact aspects of manipulative skill which can only be assessed by direct observation. If periodic grading is used, it is important that teachers should consciously look for evidence of the skills throughout the period and not simply rely on their memory at the end of the period.

Example B. Skill in observation and accurate recording of observations.

10-9 For observational exercises: correct observations from reagents specified, unexpected results noted. For quantitative exercises: errors and inaccuracies not ignored, results lying outside the range noted. Recordings of results: all relevant information accurately recorded in a form appropriate for calculation or evaluation.

8-7 Good presentation of data but little attention to errors, such as a result outside the normal range. For observational exercises: largely complete description but lacking in finer detail.

6-5 Adequate presentation of data. Inclusion in mean of wide spread of values. Some essential features omitted from description.

4-3 Poor presentation. Students' attention had to be drawn to point they should have seen for themselves. Many ommissions of essential features in description.

2-1 Very poor presentation, only part of data reported. Result/observation not acceptable. Relies upon other students.

<u>Observation</u> alone has little merit unless it is accurate and relevant. Its assessment leans heavily on a student's ability to report, and it is inevitable that most of it will take the form of awarding marks for written records of observation. But there are two aspects of observation for which teachers can give credit while a conventional practical examination cannot, and it is hoped that neither will be overlooked: (1) The first is oral reporting which is usually more immediate and less rehearsed than written reporting. (2) The second is a student's observation and interest in unusual and unexpected features of an experiment even though they are not relevant to the immediate purpose of the experiment.

Appendix 4

Examples of Report Schedules for Continuous Assessment

One example of a partially completed report card is shown below:*

NAME ______________________ TEACHER ______________________

GRADE ______________________ ROOM ______________________

SUBJECT ____

Exercises used for assessment (name or brief description)	Date	Marks awarded to qualities tested				
		A	B	C	D	E

* Adapted from JMB (1979)

TEACHER'S ASSESSMENT SCHEDULE*

DATE____________________ NAME OF TEACHER____________________

SCHOOL____________________ CLASS____________________

NAME OF EXPERIMENT____________________

NAME OF STUDENTS EVALUATED (1)____________________

(2)____________________

(3)____________________

(4)____________________

OBSERVATIONS

A. Conducting observations
B. Reporting observations
C. Confidence in observing

1	2	3	4

PLANNING

A. Design
B. Sources for errors
C. Understanding the goals of the experiment
D. Understanding the connection between the results of the experiment and goals

1	2	3	4

MANIPULATIVE SKILLS

A. Correct handling of apparatus
B. Correct sequencing of tasks
C. Effective and purposeful utilization of equipment
D. Precision of work

1	2	3	4

Manual Dexterity

1	2	3	4

Conclusions and Generalizations

1	2	3	4

* Cohen, (1977).

ERIC