

# On Evaluation

Mustafa Hassan Badi, MBBCh, FRCS, PhD\*

Socrates asked his pupil Plato; 'What is knowledge?'. Plato replied; 'to know is to perceive something with the senses'. Even before that Aristotle was also very concerned with the relation between 'knowledge' in its own essence and knowledge of individual images of things. He was also concerned with the problem of consciousness, self, and cognition; and the role of senses and reason play in the process. These concepts of knowledge have since early history and until today remained basically a philosophical problem.

A philosophical attempt to define knowledge which may appeal to some educationalists may be found in the statement of Mikbailov: "to say that I know implies the ability to look at one's own activity from the side, to consider the object before one and what it will become when one does this or that with it, .... when I can treat my activity as something ahead of me, as a future process that I can adjust or change in accordance with a pre-arranged and not yet executed ideal plan of an action, only then can I and should I consider the question of the ability to know". On the other hand the definition of knowledge can also be extracted from the old Chinese proverb "I hear, I forget; I see, I remember; I do, I know".

The above concepts were presented to illustrate: "what is knowledge"? was a primary concern of man for more than 2000 years and remained an abstract quality of mind that has no intrinsic properties in its own right. What we consider as 'knowledge' is in reality the outward manifestations of what ever underlying mental process takes place. In other words man's behavioural manifestations of knowledge.

The difficulties we have with our methods of evaluation, assessment and examinations are probably due to our inability to treat and distinguish between "knowledge" and the "manifestations of knowledge" as two separate entities.

Knowledge manifests itself as simple, complex or highly complex behavioural patterns which are theoretically evaluable. They vary from simple factual recall or recognition to conceptual level where facts are related to each other in some form of pattern, intellectual functions,

manual skills and finally as attitudes where a value is attached to the behavioural patterns themselves.

In our system of education, from primary to higher education, a great value has been attached to the ability to memorise and recall information in an examination or life situation. None or very little consideration is given to the ability to use what is learnt in a useful and observable physical activity or as an intellectual or cultural make-up worthy of value.

The subject matter, a text book, a lecture etc. have become the objects of learning and an end in themselves. For example, the text book of Anatomy is learnt, and not the Anatomy of human body. Medical graduates know all the information about the human body and are able to cough it out on request. But when manually manipulating the real structures in the living body it is soon realised that knowledge is not what is believed to have been. On the other hand, reading, seeing or listening to the description about how an injection is done does not necessarily mean its perfection in practice.

Prophet Mohammed, Peace Be Upon Him, Said; "O Allah protect me from that knowledge which is useless".

The objectives, contents, method of instruction and evaluation of educational course can address themselves to any or a combination of the outward alluded to manifestations of knowledge. These include: (a) Cognition level of recall and recognition, ie. the factual level of knowledge, (b) Level of skills and conceptual aspects of problem solving, ie, the physical manifestations of knowledge and (c) The level of attitudes.

To approach the ideal, many innovations in the area of evaluation have been introduced eg. computers, paper simulation, etc. However, there remain several widely used methods and these include: Multiple Choice Questions (MCQ), short and long essays, practical tests, oral and clinical examinations.

None of these however is fully satisfactory when compared in accordance with the accepted criteria of a good method for the evaluation of student's achievements

---

\* Vice Dean and  
Professor of Anatomy & Surgery  
Faculty of Medicine  
Um Dorman Islamic University  
Sudan



and competence. Some of these criteria of a reasonable method of evaluation are:

1. *Validity* - The examination must be a test of what to purports to assessment, eg. an MCQ or essay question cannot test the ability of the student to examine a patient.
2. *Reliability* - The examination must be used as an accurate measuring devise for testing so that must give different examiners the same conclusion.
3. *Objectivity* - The examination is a test of whether the objectives of the course are achieved or not.
4. *Feasibility* - The examination can be conducted within the physical limitations and constraints.

When applying the above criteria to the types of criteria mentioned earlier several understandings can be established:

1. MCQ examinations are reliable and feasible but tend to examine low levels of knowledge. Problems inherent in certain disciplines make it difficult to design good MCQ questions. The main difficulty resides in furnishing good alternatives to the correct answer that are not trivial. Bad MCQs on the other hand encourage bad learning habits in the students who become concerned with trivial details of the subject matter rather than broader understanding of the field of knowledge.
2. Long essay questions have a very poor inter-examiner reliability but are of greater inherent validity, and are both feasible and objective. They can test conceptual knowledge and the student's problem solving abilities. They are very much affected by the student mastership of language which is a major problem in many schools.
3. Short essays have a higher inter-examiner reliability than long essays, a greater validity than MCQs and are feasible. They may, however, demonstrate factual knowledge only. Their demand on language is minimal.
4. Practical examinations which can test manual and

conceptual skills may be the best form of examinations available but must be thoroughly and objectively designed.

5. Oral examinations can be valid and feasible but they are unreliable and un-reproducible.

Therefore, the dissatisfaction with the methods of examination should suggest other measures for assessment such as:

1. Use different forms of examinations such as MCQ, short and long essays, practical tests, oral and clinical examinations.
2. A large number of examiners correcting the same paper or assessing the same student increase inter-examiner reliability.
3. Use a closed system of marking in every single question and in the examination as a whole. This again increases inter-examiner reliability and reduces the gravity of a single examiner judgement. The marking system may be fixed at a range between 45% -75% with a pass mark of 50%.
4. Accepting a failure level of 49% as marginal which can be increased through compensation by in-course performance. Marks from other parts of the same examination, or from examinations in other related subjects provide a third checking mechanism.
5. The examiners must function collectively and not individually and are expected to justify their judgement to examiners and college boards as well as to the University Senate before any results are considered final.

In conclusion, there remained many aspects that need to be investigated and achieved in the area of evaluation. There is a serious concern and an awareness of the need for innovations and change in the methods of evaluation, and attempts must be made to find a reliable examination method which can test the worthy qualities of students.

### Further Reading

Mikbailov. The Riddle of the self. Moscow: Progress Publishers, 1980: 25-82.