

FORMATIVE CLINICAL EVALUATION OF FIRST-YEAR STUDENTS IN FUNDAMENTAL NURSING SCIENCE

I. Wannenburg, W. J. Kotzé

"The purpose of evaluation is not to prove but to improve"
(Morgan & Irby, 1978, p. 176)

The outcome of any system of evaluation has a considerable impact on the daily lives of students in a variety of educational settings. Hence, it was attempted to illustrate the complexity of the problem when evaluating the performance skills of first year students in the clinical environment of the general hospital.

The choice of the research field originated from the increasing concern of the researcher about the reliability of the current evaluation practices in the formative assessment of first year student nurses. The impression gained is that nurse educators are more concerned with the end results of evaluation than with the teaching-learning process needed to reach this goal. Due to the many variables that can influence its results, the implementation of the evaluation process in clinical nursing is extremely complicated. In the course of studying the literature relevant to the research field, the researcher identified aspects that can be considered as of critical importance in the assessment of student performance in clinical nursing.

ESSENTIAL REQUIREMENTS FOR FORMATIVE EVALUATION

Evaluation criteria

The following evaluation criteria have been identified as essential for both formative and summative evaluation:

- criteria to measure the reliability of the measuring instrument
- criteria to ensure safe patient care
- criteria to facilitate the integration of theory and practice
- criteria to assess student learning whilst mastering nursing skills

Variables

The variables with direct influence on the reliability of the outcome of evaluation have been found to be the following:

- Human behaviour, in principle is difficult to observe and evaluate, so that some degree of error should always be anticipated
- Students' clinical experiences are obtained in a variety of clinical situations, so that a uniform system of individual student evaluations are

complicated by factors such as a mix of patients, the dissimilarity of their illnesses, personalities and nursing needs

- Evaluators and students, being part of a pluralistic society, will invariably interpret data according to their own culture, standards and values
- Evaluation of a student's performance in the clinical environment is done by direct observation, which necessitates the setting of behavioural objectives to serve as criteria for measurement, irrespective of the measuring instrument utilized.

Safety standards

The inclusion of a patient in the teaching-learning milieu demands the inclusion of safety standards for the performance of nursing skills.

Integration of theory and practice

Integration of theory and practice does not occur automatically. This may be facilitated by the nature of planning and utilization of learning content and opportunities in the learning environment, as well as by the integration of cognitive learning in the criteria of clinical measuring instruments.

Progress of student learning

Formative evaluation focusses on the attainment of teaching and learning goals. Criteria, therefore, which reflects both development in skill learning and synthesis of knowledge, should be a requirement in all clinical performance measuring instruments.

Because Nursing Science is defined and accepted as a clinical health science, all the above-mentioned aspects are regarded to be of vital importance in the training of student nurses. The assessment of performance in the clinical field must be a reliable reflection of a student's achievement, as the outcome of such evaluations will have a considerable influence on the student's position in the training programme and his/her future career prospects. It will contribute towards determining whether a student:

- is progressing towards reaching nursing programme, curriculum, stage and learning objectives
- has met the requirements for admission into an examination
- may be promoted to the following stage

- or year of study
- qualifies for awards or bursaries
- has complied with all the programme objectives required for certification of completion of training, in order to be recognised by the South African Nursing Council for registration purposes.

RESEARCH APPROACH

Data collection and research instrument

A descriptive statistical approach was implemented. The target population consisted of the twenty-two (22) South African Nursing Council approved nursing schools which were selected as appropriate for the study.

The data for the research was acquired by obtaining the relevant formative measuring instruments from all nursing schools after having received permission from the relevant controlling health authorities.

For the assessment of the evaluation instruments a reliable approach had to be found. The researcher chose the approach of Battenfield as the appropriate one for this study because it encompasses all the criteria identified in the literature as of critical importance in formative clinical nursing evaluation. A checklist containing this author's fourteen criteria was compiled.

In the first phase of the analysis the macro design of every evaluation instrument received from the research population was assessed by using these fourteen criteria. In the analysis each of the criteria was regarded as of equal importance.

Following this the instruments were assessed for their compliance with the requirements of the four major critical formative criteria, namely:

- to measure the reliability of the measuring instrument
- to ensure safe patient care
- to facilitate the integration of theory and practice
- to assess student learning whilst mastering nursing skills.

The grouping of the fourteen criteria of Battenfield into the above four major critical areas was done by the researcher according to priorities identified in the literature studied. It was done in this way because no proof could be found of any other, more scientifically based, e.g. through statistical analyses, approach. Based on her experience in this regard the

researcher recommends that a comprehensive study be done in which for example, factor analysis could be utilized to verify the most critical nursing areas for evaluation purposes.

Formulation of a hypothesis

The hypothesis formulated for the research was as follows: "The formative clinical measuring instruments utilized currently to assess first year student nurses in the clinical environment of the general hospital do not promote student learning".

Data analysis and recommendations

It must be emphasized that the whole research population came under discussion. The response from the relevant nursing schools was 86,3%.

Number of clinical measuring instruments in use

The total number of clinical measuring instruments obtained from the nursing schools was 254 and varied between 1 to 62 per school (compare Figure 1). Only 36,8% of the nursing schools submitted ten or more clinical measuring instruments to assess performance skills. The criterion-referenced grading system was utilized by 73,7% of nursing schools. A checklist was the selected instrument utilized to assess the first year student nurse's performance skills in the clinical milieu.

A checklist basically records whether a characteristic is present or absent or whether an action was taken or not, it requires a simple "yes-no" judgement. Checklists are useful when evaluating those performance skills that can be divided into a series of clearly defined, specific actions where appraisal is so rough that one is limited to only two options.

Recommendations

The emphasis in formative evaluation should be on guiding the student towards

self-directed learning. By providing regular feedback, using clear, precise and measurable evaluation criteria, the student is encouraged to become more self-evaluative and actively involved in the learning process. The art of evaluation lies in not only stating that something is wrong, but indicating why it is wrong and, most important, which measures should be taken to correct the nursing actions. It is at this stage that clearly stated objectives and measurable evaluation criteria support the evaluator in guiding the student constructively. A specific measuring instrument should be designed for every identified nursing skill in which the student has to prove her competency. This approach to formative evaluation will not only increase the reliability of the measuring instrument but also enhances learning. It can, therefore, be utilized by both the tutor and the student to improve teaching and learning skills.

Incidence of individual evaluation criteria

The frequency of the evaluation criteria found in the 254 obtained measuring instruments were as follows (compare Figure 2).

Criterion 1: Frequency 1

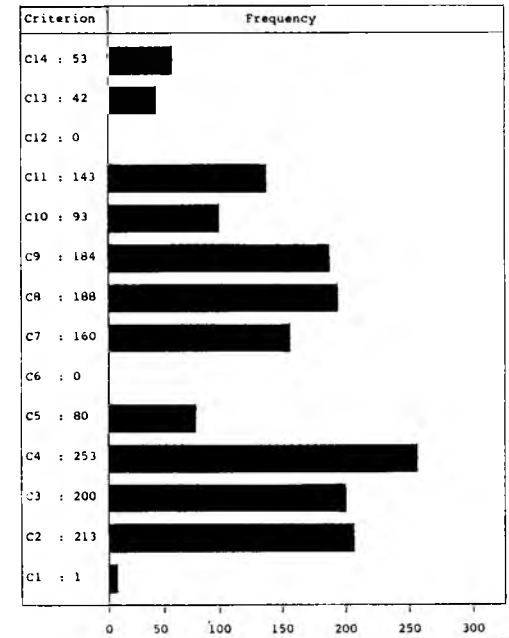
Does the measuring instrument include clinical objectives that reflect the curriculum and stage objectives?

This criterion was only reflected in one of the measuring instruments, namely 74. This could possibly be a reason for the problems that exist with theory and practice integration in nursing education. The above finding and statement supports the concern found in the literature in this regard.

Recommendation

It is suggested that programme, curriculum, classroom and stage objectives be included in the performance measuring

FIGURE 2: FREQUENCY OF USE OF EVALUATION CRITERIA



instruments to facilitate the integration of theory and practice.

Criterion 2: Frequency 213

Does the measuring instrument contain items that are attainable, clear and stated in behavioural terminology?

This criterion was met by 83,9% of the presented measuring instruments, only 16% did not meet this criterion.

Recommendation

It can be deduced from the above findings that most nursing schools are utilizing this criterion in the clinical measuring instruments.

Subjectively and vague expectations can be eliminated by stating precisely what is required, for example, "all" equipment must be itemized, "correct" records must be specified, and so on.

Starting the statements with a verb and concentrating on performance actions will help the nurse educator to state test items that are attainable and easier to assess, for example: "Administers medications on time".

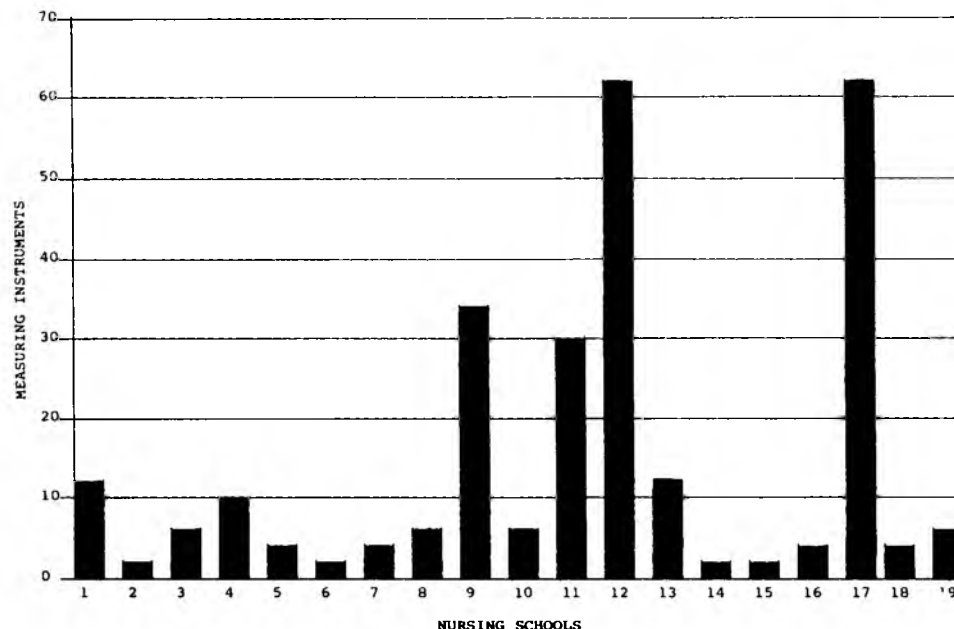
Criterion 3: Frequency 200

Does the measuring instrument contain items that describe behaviours that are measurable?

This criterion is very closely associated with the previous criterion, because if the test item is not stated in behavioural terminology it cannot be measured, 78,7% of the measuring instruments met this criterion.

The following was observed in the 21,3% of measuring instruments that did not meet this criterion. It is, for example, impossible to measure abstract behaviour such as, "professional courtesy and compassion through verbal and non-verbal communication", "displays responsibility

FIGURE 1: NUMBER OF MEASURING INSTRUMENTS OBTAINED FROM EACH NURSING SCHOOL



with regard to the level of understanding and language ability of the patient”, and “demonstrates knowledge and insight”.

Recommendation

The more specific the wording of the individual test items the more reliable the results. A verb that is open to varied interpretations must be avoided. If it is used it must be accompanied by a description of the exact expected behaviour. Verbs such as know or understand are difficult to evaluate and should be excluded except if they are supplemented by action statements. Other examples of terms that should be avoided are those that are not observable, for example, enjoy, believe, grasp the significance of, respect, always and never (compare Battenfield, 1986, p. 19).

Criterion 4: Frequency 254

Does the measuring instrument appear appropriate for the clinical setting in which it is used?

The highest frequency is obtained by this criterion which reflects the appropriateness of the designed measuring instruments for the hospital as clinical evaluation milieu.

No recommendation is required for this criterion.

Criterion 5: Frequency 80

Does the measuring instrument relate classroom objectives to the performance skills?

As indicated by the low frequency obtained, namely 31,5%, it is obvious that this criterion has not been achieved.

Recommendation

It is recommended that classroom objectives be included in performance measuring instruments as it not only enhances theory and practice integration but also excludes the danger of performing any psychomotor skill without the necessary cognitive basis.

Criterion 6: Frequency 0

Does the measuring instrument demonstrate progression of skills and synthesis of learning?

Not one measuring instrument demonstrated a progression of skills or a synthesis of learning.

Recommendation

Formative evaluation focusses on student learning. If criteria are not identified to assess progression of skills and synthesis of learning then the evaluation in this area is subjective. It can also be deduced that higher level skill learning is not assessed in the clinical environment. It is, therefore, recommended that clinical performance measuring instruments include criteria that would make it possible to assess high level skill learning.

Criterion 7: Frequency 160

Does the measuring instrument contain clearly written descriptors for the criteria used in measuring student achievement?

This criterion was met by 63,0% of measuring instruments. Confusion still exists in 37,0% of the presented measuring instruments.

The following are not clearly described:

- outstanding clinical performance
- satisfactory clinical performance
- unsatisfactory clinical performance

In addition to this, terminology such as competency and proficiency are used interchangeably. A key legend is absent for describing numerical values and terminology such as occasionally, seldom and frequently.

Recommendations

Clearly descriptive terminology provides for reliable and fair student assessments. It is suggested that the following terminology be used. Compare Battenfield (1986):

“**Outstanding.** Consistently superior in clinical performance, skill, synthesis of learning and application of nursing principles. Functions above expectations of student at this level.

- Is able to function safely and effectively with minimal guidance.
- Demonstrates superior ability to make alert and informed observations.
- Demonstrates superior ability to synthesize knowledge, clinical performance is beyond expected level.
- Seeks out and assumes responsibility beyond that required.
- Makes appropriate nursing judgements with minimal guidance.
- Initiates appropriate nursing intervention with minimal guidance.
- Consistently seeks learning experiences.
- Consistently demonstrates expected professional conduct.

Satisfactory. Safe clinical performance; demonstrates expected skills, synthesis of learning and application of nursing principles at expected level.

- Functions safely with guidance.
- Demonstrates ability to make expected observations.
- Demonstrates ability to relate and apply knowledge; clinical performance at expected level.
- Assumes responsibility for assignments.
- Makes appropriate nursing judgements with guidance.
- Initiates appropriate nursing interventions with guidance.
- Usually seeks learning experiences.
- Adequately demonstrates expected professional conduct.

Unsatisfactory. Clinical performance inadequate; indicates lack of skill, unsafe nursing practice, inadequate depth of knowledge, or application of nursing principles. Functions below expectations of student at this level.

- Does not function safely even with guidance.
- Does not demonstrate ability to make expected, obvious observations related to patient care.
- Does not demonstrate ability to relate

knowledge to clinical situation at expected level.

- Has difficulty accepting responsibility.
- Does not make safe nursing judgements.
- Does not initiate nursing intervention without maximum guidance.
- Seldom seeks learning experiences.
- Demonstrates unprofessional clinical conduct.”

(Battenfield, 1986, p. 27-28)

Terminology such as “occasionally” and “seldom”, for example, should be supported with clear descriptions of the expected behaviour.

Criterion 8: Frequency 188

Does the measuring instrument identify the critical elements, if any, that have to be met?

This criterion was met by 74,0% of nursing schools. These nursing schools also clearly indicated in their instructions to the evaluator how these critical elements should be assessed in the final calculation for a pass-fail grade. However, 26,0% of nursing schools are not sure of the meaning of a critical element or how to distinguish it from steps that have to be followed to learn a skill.

Recommendation

With the addition of a patient into the teaching-learning situation the inclusion of safety standards for patient care should receive top priority. This is obtained by identifying the critical elements for each nursing skill. These statements do not tell how the task should be accomplished but focus on the behaviours that are essential for the patient's protection. These critical elements should when finalized, explicitly state the essential nursing behaviours that constitute acceptable practice and heed 100 percent accuracy to ensure safe patient care. The emphasis is not on the needs of the student but on the student's behaviour in relation to the patient's needs.

Critical elements are thus the observable and mandatory behaviours that collectively comprise the standard against which the student's competence is measured. When one of them is violated or omitted the patient is actually or potentially endangered and nursing care is unsafe. The student therefore, fails the performance evaluation no matter how well she has performed in any of the other criteria (Lenburg, 1979, p. 47-48).

Criterion 9: Frequency 184

Does the measuring instrument clearly describe the grading scale for determining a pass-fail grade?

This criterion was achieved by 72,4% of measuring instruments. A final calculation for determining a pass-fail grade is obtained by integrating the mark allocated for the theoretical and the performance evaluations. Confusion exists when, in the final evaluation, the theoretical knowledge is assessed by using a normative-referenced grading scale and the performance

evaluation is assessed by using a criterion-referenced grading scale. Calculating this final pass-fail grade then becomes a very difficult task.

Another problem encountered is complicated weighting of numerical systems, for example: that of 3 test items to which 5 marks have to be allocated. As the evaluator receives no instruction on how the marks are to be divided between the items, the nature of the calculation is left to the discretion of the individual evaluator, resulting in subjectivity and thus, unreliability.

Recommendation

When formulating a final judgement on the student's clinical performance the clinical measuring instrument should describe the grading scale for determining a pass-fail grade. It is suggested that the following be used as a guide:

- identification of the total number of clinical objectives that must be obtained
- compliance with critical elements
- provision of a key for assigning numerical values and a formula for calculating numerical ratings (compare Battenfield, 1986, p. 29).

Since comparative grading is not an issue in a criterion-referenced system, the student must pass, or demonstrate competency in every required skill. The student cannot compensate for failure on one skill by passing another. This problem is encountered when a numerical grading system is used. A student may, for example, obtain 80% for measuring a patient's blood pressure, 50% for bed bathing a patient and 20% for urinalysis. The average mark for all these skills may then be calculated as 50%, but 50% of what?

Complicated weighting and numerical systems should be avoided as this is confusing and reduces the reliability of the instrument.

Criterion 10: Frequency 93

Does the measuring instrument provide a column for notation of not applicable or not observed behaviour?

Only 36,6% of measuring instruments made provision for this criterion.

Recommendation

A student cannot be judged on a clinical objective or on an individual test item not applicable to a situation or which the evaluator failed to observe. It is unfair and results in unreliable evaluation.

Criterion 11: Frequency 143

Does the measuring instrument provide the evaluator with a space for writing supporting statements and, if needed, a recommended plan for remediation?

Only 56,3% of measuring instruments reflected this criterion.

Recommendation

All measuring instruments should reflect this criterion as formative evaluation focusses upon teaching-learning strategies. The evaluator, therefore, needs a space on the measuring instrument for writing supporting statements, comments and, if needed, a recommended plan for remediation.

Criterion 12: Frequency 0

Does the measuring instrument promote interrater reliability by providing a guide for writing supportive statements that are congruent with the evaluation rating scale?

The incidence of this criterion in the measuring instruments was zero.

Recommendation

It is suggested that this criterion be included, as the reliability of the measuring instrument will be enhanced by the writing of supporting statements, comments and, if needed, a recommended plan for remediation (compare Battenfield, 1986, p. 38).

Criterion 13: Frequency 42

Does the measuring instrument provide students with a space for writing optional statements?

The incidence of this criterion was remarkably low, namely 16,5%.

Recommendation

The student should be actively involved in the evaluation process and has the right to approve of, or disagree with what the evaluator has recorded, in writing on the measuring instrument.

Criterion 14: Frequency 53

Does the measuring instrument provide space for signature and date lines for both the evaluator and the student?

This criterion was met by 20,9% of the measuring instruments. Most measuring instruments made provision for the

signature and date lines for the evaluator but not for the student. This fact once again caused the researcher to question how actively the student is involved in her own learning process, or whether this information remains in the possession of the evaluator only.

Recommendation

If the student is to be actively involved in her own learning process she should also be actively involved in the evaluation process. The instrument should, to identify this, therefore also provide space for signature and date lines for student use.

Ratings obtained for reliability, safe patient care, facilitating theory and practice integration and progress in student learning

When the instruments were assessed for compliance of the criteria with: reliability, safe patient care, facilitating integrating of theory and practice and progress in student learning, the following was found:

A measure of the *reliability* as being measured by these five criteria, may be taken as:

$$\frac{\text{sum of the frequencies of the five criteria}}{\text{number of criteria in group}} \times 100\%$$

$$= \frac{759}{5 \times 254} \times 100\%$$

$$= 59,8\%$$

The second highest rating 59,8% was obtained for the criterion *reliability*. The reliability of measuring instruments can be increased by promoting interrater reliability, describing rating keys, formulating behavioural objectives that are clear, precise and measurable, and constructing a measuring instrument for each nursing skill that has to be assessed.

The more specific the wording of the test items, rating items and instructions to evaluators, the more reliable the results. By minimizing the number of judgements required, for example by using "yes/no" or "satisfactory/unsatisfactory" response options, the more likely it will be that

FIGURE 3: RELIABILITY

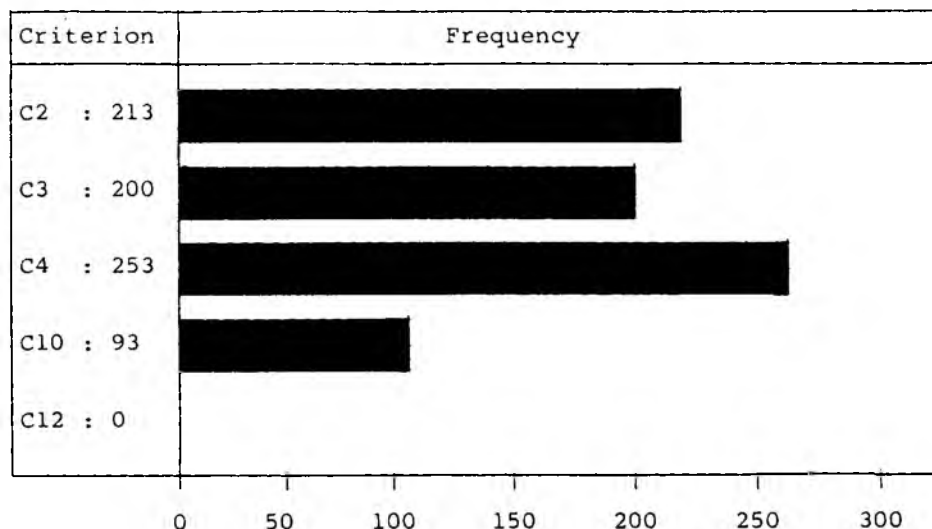
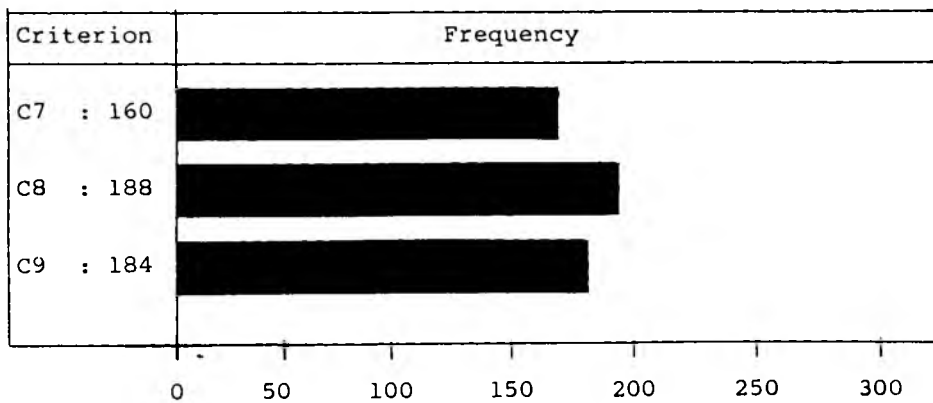


FIGURE 4: SAFE PATIENT CARE



agreement among raters and, therefore, higher interrater reliability will be achieved.

The evaluator may want to elaborate on the observed behaviour by writing comments or supportive statements. These recordings may be to highlight positive behaviour or to clarify statements of negative ratings.

A list of descriptors to guide supplementary documentation of the observed behaviours can further enhance interrater reliability. For greater reliability adequate sampling of experiences, representing all desired behaviours, is preferred. With the variables inherent in clinical situations it is necessary to assess the student in a range of learning situations to ensure reliability of judgement. No student should be judged

on a clinical objective unless the evaluator has actually observed the behaviour. Guessing is inaccurate and subjective. To allocate a zero in an unfilled rating box in such an event is unfair to the student. Therefore, it is necessary to provide a means for notation of "not applicable" or "not observed" behaviour.

A measure of the *safe patient care* as being measured by these three criteria, may be taken as:

$$\frac{\text{sum of the frequencies of the three criteria}}{\text{number of criteria in group} \times 254} \times 100\%$$

$$= \frac{532}{3 \times 254} \times 100\%$$

$$= 69,8\%$$

FIGURE 5: FACILITATING THEORY AND PRACTICE INTEGRATION

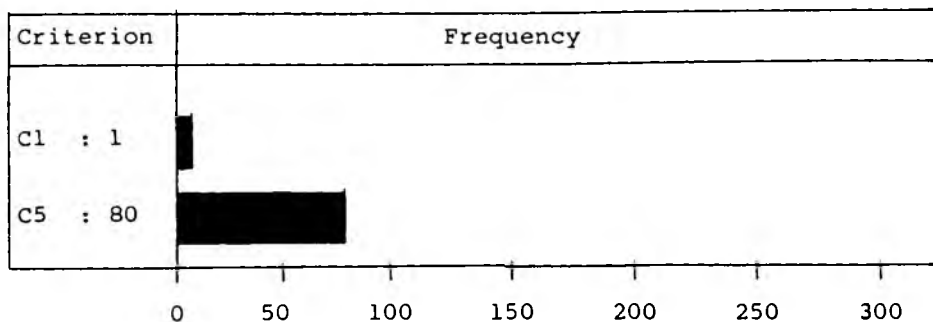
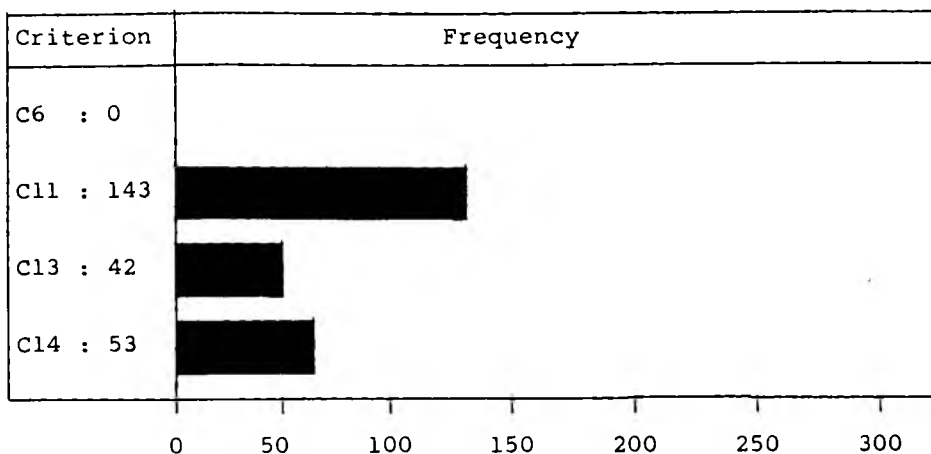


FIGURE 6: PROGRESS IN STUDENT LEARNING



Safe patient care received the highest rating 69,8%, which indicates the concern of nurse educators for patient safety. It must in addition be emphasized that this is also the most important criterion in summative evaluation. Considering the low rating obtained by the other evaluation criteria, student learning is not the main concern in the formative evaluation milieu.

A measure of the *facilitation of theory and practice integration* as being measured by these two criteria, may be taken as:

$$\frac{\text{sum of the frequencies of the two criteria}}{\text{number of criteria in group} \times 254} \times 100\%$$

$$= \frac{81}{2 \times 254} \times 100\%$$

$$= 15,9\%$$

Great concern has been expressed about the lack of integration between the theory and practice in nursing. The low incidence, 15,9%, in this study supports the concern expressed in the literature. The nurse educator plays an important role in facilitating the integration of theory and practice by structuring, organizing and presenting nursing curricula in such a manner that retention and transfer of knowledge can occur in the student. This can be enhanced by including clinical objectives in the formative measuring instruments that reflect the curriculum, stage and classroom objectives.

A measure of the *progress in student learning* as being measured by these four criteria, may be taken as:

$$\frac{\text{sum of the frequencies of the four criteria}}{\text{number of criteria in group} \times 254} \times 100\%$$

$$= \frac{238}{4 \times 254} \times 100\%$$

$$= 23,4\%$$

Progress in student learning. Criteria used in the general measuring instruments submitted for this study were so non-specific that they cannot promote student learning. No measuring instrument demonstrated criteria assessing the progression of skill and synthesis of learning. It can, therefore, be deduced from a rating of 23,4% that the focus in formative clinical evaluation is not student learning.

Support for the stated hypothesis

It can be deduced from the above findings that the current formative clinical measuring instruments used by training schools included in this study to assess the performance skills of the first year basic nursing students in the clinical environment of the general hospital, do not promote student learning in all its facets. The stated hypothesis, namely, that "The formative clinical measuring instruments utilized currently to assess first year student nurses in the clinical environment of the general hospital do not promote student learning", is therefore supported by the afore discussed research outcome.

Study deficiencies

As the individual nursing schools are still in the process of developing measuring instruments to assess the individual performance skills, a representative sample of these measuring instruments could not be obtained to present any reliable judgement on the micro design. It would therefore be impractical to design a measuring instrument for any one skill tested, as one would require a response of at least 80% from the research population in each case before any valid or reliable instrument could be drawn up, based on the research findings.

The researcher failed to obtain the following information from the nursing schools:

- whether the assessment is documented in the student's presence
- whether the student is supplied with a copy of the feedback
- the role of the tutor in structuring, organizing and presenting nursing curricula in order to facilitate theory and practice integration
- the method of dealing with a student's inadequate clinical performances
- the number of performance evaluations done per student
- the nursing skills which are regarded as critical for clinical nursing assessment in the first year of study.

FINAL REMARKS

The researcher has come to the conclusion that the nursing education fraternity is to a large extent guilty of rushing through the process of evaluation — more concerned about its completion than its substance. Most of the problems encountered can be

solved by thorough design of the measuring instruments. It must be taken into consideration that nursing schools are currently in the process of not only implementing a totally innovative basic nursing programme, but also of developing clinical evaluation measuring instruments to appraise the outcome of these programmes.

Evaluation, as an essential part of any educational programme, needs to be continuous in a meaningful criterion-referenced instruction. If a student does not progress as expected, feedback should help to locate the source of reason for failure. It may be due to failure by the student, the tutor may be at fault for using ineffective teaching methods or creating inadequate learning opportunities and thus causing problems related to teaching-learning strategies.

The lack of appropriate measuring instruments for evaluation results in clinical nursing performance assessments that are often subjective, vague, unreliable and meaningless. Until accurate and reliable assessment techniques to measure students' progress towards all learning objectives have been pursued, evaluation cannot be accepted and offered as a scientific process.

Evaluation acts as a tool to ensure that the standards of nursing care are maintained and improved where necessary. The protection of the patient, student and tutor is, therefore, dependent upon the utilization of reliable assessment techniques and approaches in the clinical environment.

The nursing school is ultimately accountable to the South African Nursing Council for the standard of its curricula

and the quality of nurse practitioners it supplies to society. The South African Nursing Council has the power to inspect these schools at regular intervals to ensure that students are trained at a professional and personal level and that the needs for health care of the consumer of the South African society are met.

In 1978 the National League for Nursing expressed concern regarding the subject of evaluation, "Making decisions about anything in life is a complex and sometimes frustrating experience. Somewhere deep in the regions of our psyche lingers a fear that we might make the wrong decision. This is most true when we are dealing with decisions that affect other human beings. Another unfortunate characteristic in decision making is that enough valid data are seldom available to assist in the process" (National League for Nursing, 1978, p. 63).

The researcher believes that the fear of subjective and unreliable outcomes in the utilization of clinical evaluation instruments will in future be eliminated only to the extent that nurse educators succeed in introducing more scientifically based evaluation practices.

BIBLIOGRAPHY

- Battenfield, B.L. 1986. *Designing clinical evaluation tools: the state of the art*. New York: National League for Nursing, 43p.
- Morgan, M.K., Irby, D.M. 1978. *Evaluating clinical competence in the health professions*. Saint Louis: C.V. Mosby company, 316p.
- National League for Nursing. 1978. *Concepts and components of effective teaching*. New York: National League for Nursing. 63p.