# Practices and Challenges of Assessing Graduate Students Learning Outcomes in Ethiopian Higher Education Institutions

*Desalegn Chalchisa Jebena (PhD)*

Hiroshima University Visiting professor

July 2013

# Outline of the presentation

1. The problem  of the study
2. Methods of the study
3. Data presentation and analysis
4. Summary and recommendations

# 1. The Research Problem

- the purpose of the study is to explore the practices and challenges of instructors' in the assessment of students learning outcomes.
- The study investigates assessment of learning outcomes in terms of:
  - types and frequency of assessment strategies,
  - types of test item in practice in HEIs,
  - use of assessment results for improving learning.

# 1. The Research Problem …

## Objectives of the study

- The general objective of the study is to explore the practices and challenges of assessing student learning outcomes in selected HEIs in Ethiopia.
- The specific objectives of the study are to:
  - determine the assessment strategies practiced by instructors in assessing the learning outcome of students in HEIs;
  - identify the type of test items instructors use in assessing graduate students learning outcomes in HEIs;
  - explore how instructors use results of assessment of graduate students learning outcomes to improves students learning and instruction;
  - analyze the alignment of the methods of students' assessment with the learning outcomes;
  - Evaluate the appropriateness of the methods of students' assessment in HEIs and thereby to suggest the effective methods for better students learning.

# 2. Methods of the study

## 2.1. Design of the study

- Qualitative and quantitative approach
- The concurrent QUAN and qual design
- The descriptive design specifically survey method
- Dominantly quantitative data using questionnaire
- Qualitative data using interviews and open ended questions to complement the quantitative data.

# 2. Methods of the study …

## 2.2. Population, sampling and sample sizes

Purposive and stratified sampling were used

| University | Band/Field of study | Total number of instructors* | | Sample size | | Total number of students* | | Sample size | |
|---|---|---|---|---|---|---|---|---|---|
| | | F | M | F | M | F | M | F | Male |
| Adama | Engineering and Technology | 15 | 356 | 2 | 31 | 77 | 697 | 7 | 68 |
| Addis Ababa | Social Sciences and Humanities | 50 | 272 | 6 | 38 | 783 | 3137 | 75 | 302 |
| Bahir Dar | Natural and computational Sciences | 122 | 982 | 1 | 15 | 44 | 456 | 4 | 44 |
| Hawassa | Agricultural and life Sciences | 15 | 124 | 2 | 11 | 39 | 194 | 4 | 18 |
| Jimma | Medicine and health sciences | 45 | 297 | 4 | 26 | 63 | 495 | 6 | 48 |
| Mekelle | Business and Economics | 22 | 139 | 2 | 12 | 21 | 223 | 2 | 22 |
| **Total** | | **269** | **2170** | **17** | **133** | **1027** | **5202** | **99** | **501** |

# 2. Methods of the study …

## 2.3. Instruments

- Student questionnaire,
- Teachers' questionnaire and
- Interview guides
- The questionnaires were adopted from Hale and Astolfi (2011) of Strategies Used for Assessing Student Learning Survey.
- Similar items were included in the questionnaire for teachers and students.
- The instruments measured:
  - Characteristics of respondents
  - Types of assessment strategies and their frequent uses
  - Types of test items and their frequent uses
  - Use of the result of assessment for improving student learning.
- Each subsection contained items with a five point scale of Never (1), Rarely (2), Often (3), Frequently (4), and Very Frequently (5).

# 2. Methods of the study …

## 2.4. Procedures

- The scales of Hale and Astolfi (2011) for surveying assessment strategies was used with modifications.
- The interview guides were prepared to complement data collected through questionnaire. Comments on the items were collected from PhD students of the curriculum instruction to ensure the content validity of the questionnaire and interview guides
- Questionnaire was pilot-tested on instructors and students that were not participated in the final study
- The Cronbach alpha indices of reliability of the instruments were 0.86 for instructor and 0.91 for student questionnaires.
- For each sub-scores, scaled items were converted to T-scores for analysis.

# 2. Methods of the study …

**2.5. Methods of data analysis**

- Percentages were used to find out types and frequencies of assessment strategies, types and frequencies of test items practiced, and the extent use of assessment results for improving learning.
- Mean and standard deviations were used to describe the distribution of scores from students and instructors responses.
- One way analysis of variance (ANOVA) was used to explore differences in the mean scores of assessment strategies, types of test items and use of assessment result to improve students learning. .
- Qualitative data were organized to complement the major themes included in the quantitative analysis.

# 3. Data presentation and analysis

- The analysis was carried out under four parts:
  1. Demographic characteristics,
  2. Types and frequency of assessment strategies,
  3. Types and frequency of test items and use of assessment strategies and
  4. Use of assessment results for improving students learning.

# 3.1. Demographic characteristics

*Sex of respondents*

| sex | Students | | Instructors | |
|---|---|---|---|---|
| | N | % | N | % |
| Female | 88 | 16.8 | 5 | 3.8 |
| Male | 425 | 81.0 | 124 | 94.7 |
| NR | 12 | 2.3 | 2 | 1.5 |
| **Total** | **525** | **100.0** | **131** | **100.0** |

# 3.1. Demographic characteristics …

*Respondents by university*

| University | Bands/Field of study | Students | | Instructors | |
|---|---|---|---|---|---|
| | | N | % | N | % |
| Adama | Engineering and Technology | 62 | 11.8 | 14 | 10.7 |
| Addis Ababa | Social Sciences and Humanities | 101 | 19.2 | 36 | 27.5 |
| Bahir Dar | Natural and computational Sciences | 73 | 13.9 | 22 | 16.8 |
| Hawassa | Agricultural and life Sciences | 108 | 20.6 | 17 | 13.0 |
| Jimma | Medicine and health sciences | 88 | 16.8 | 23 | 17.6 |
| Mekelle | Business and Economics | 72 | 13.7 | 18 | 13.7 |
| NR | | 21 | 4.0 | 1 | 0.8 |
| **Total** | | 525 | 100.0 | **131** | **100.0** |

# 3.1. Demographic characteristics …

*Instructors by academic rank*

| Academic rank | N | % |
|---|---|---|
| Lecturer | 37 | 28.2 |
| Assistant Professor | 61 | 46.6 |
| Associate Professor and above | 31 | 23.7 |
| NR | 2 | 1.5 |
| **Total** | **131** | **100.0** |

# 3.1. Demographic characteristics …

*Instructors' teaching experience in years*

| Experience in years | N | % |
|---|---|---|
| 10 and below | 61 | 45.2 |
| 11-19 | 28 | 20.7 |
| 20-29 | 27 | 20.0 |
| 30 and above | 10 | 7.4 |
| NR | 9 | 6.7 |
| **Total** | **131** | **100.0** |

# 3.2. Practices of assessing students learning outcomes

## Types and frequent of learning strategies

| Item | Students | | Instructors | |
|---|---|---|---|---|
| | Highest rating | % | Highest rating | % |
| Computational problems | Never | 27.0 | Rarely | 23.0 |
| Individual project(s) | Frequently | 23.2 | Frequently | 33.3 |
| Group project(s) | Frequently | 26.5 | Frequently | 25.2 |
| Short essay | Frequently | 25.0 | Frequently | 29.8 |
| Long essay | Frequently | 23.4 | Very frequently | 21.0 |
| Individual presentations | Frequently | 25.7 | Frequently | 29.8 |
| Group presentations | Often | 27.1 | Frequently | 25.2 |
| Student lab books | Never | 44.4 | Never | 41.2 |
| Portfolios | Never | 70.1 | Never | 60.5 |
| Individual term paper | Very frequently | 33.5 | Frequently | 40.5 |
| Group term paper | Very frequently | 40.5 | Very frequently | 30.5 |

## 3.2. Practices of assessing students learning outcomes …

*Mean and standard deviation of Types of assessment strategies*

| University | Bands/fields of study | Students | | | Instructors | | |
|---|---|---|---|---|---|---|---|
| | | N | $\bar{X}$ | $s$ | N | $\bar{X}$ | $s$ |
| Adama | Engineering and Technology | 117 | 50.81 | 8.96 | 16 | 52.73 | 12.93 |
| Addis Ababa | Social Sciences and Humanities | 89 | 51.23 | 9.09 | 34 | 49.78 | 9.00 |
| Bahir Dar | Natural and computational Sciences | 70 | 47.08 | 11.49 | 20 | 46.45 | 8.17 |
| Hawassa | Agricultural and life Sciences | 87 | 47.15 | 8.69 | 17 | 47.34 | 11.02 |
| Jimma | Medicine and health sciences | 70 | 49.69 | 11.75 | 23 | 52.55 | 9.02 |
| Mekelle | Business and Economics | 71 | 53.67 | 8.87 | 18 | 51.19 | 10.42 |
| **Grand mean** | | **504** | **50.00** | **10.00** | **128** | **50.00** | **10.00** |

## 3.2. Practices of assessing students learning outcomes …

*Summary of ANOVA for strategies of assessment of learning outcomes*

| Respondents | Source of variation | SS | df | MS | F | p |
|---|---|---|---|---|---|---|
| students | Between Groups | 2476.49 | 5 | 495.30 | 5.22 | 0.00 |
| | Within Groups | 47214.29 | 498 | 94.81 | | |
| | **Total** | **49690.78** | **503** | **30.68** | | |
| Instructors | Between Groups | 668.20 | 5 | 133.64 | 1.36 | 0.25 |
| | Within Groups | 12031.80 | 122 | 98.62 | | |
| | **Total** | **12700.00** | **127.00** | | | |

# 3.2. Practices of assessing students learning outcomes …

*Types and frequency of test items for assessing learning outcomes*

| Item | Students | | Instructors | |
|---|---|---|---|---|
| | Highest rating | % | Highest rating | % |
| Multiple choice | Never | 41.5 | Never | 31.9 |
| True-False | Never | 49.5 | Never | 41.6 |
| Matching | Never | 49.2 | Never | 37.3 |
| Short Answer | Frequently | 25.7 | Frequently | 26.0 |
| Short/limited response Essay | Frequently | 27.2 | Frequently | 31.1 |
| Long/extended response Essay | Frequently | 28.8 | Frequently | 32.2 |
| Computational problems | Never | 24.1 | Never | 38.2 |
| Oral tests | Never | 56.0 | Never | 46.6 |
| Individual performance checklists | Never | 38.1 | Never | 27.4 |
| Group performance checklists | Never | 33.1 | Never | 30.6 |

# Students interview responses on types of assessment strategy

"Many instructors use only mid and final examination and correct these out of 100% and set grade on relative basis. I have seen very few of my instructors using other instruments" (GS-1)

***Comment*:** *As opposed to the HEIs assessment policies, summative evaluation is the main purpose of assessment of students in HEIs*.

# Instructors interview responses on types of assessment strategy

"In my courses, I use laboratory reports, review of scientific articles with presentation to colleagues, term papers which contribute (50%) and mid and final examinations which also contribute 50%, which would be added and the final grading be decided on norm-referenced approach."

*Comment: It seems that the instructors reflected exactly what the policy says. This completely disagrees with students responses. In agreement with this response, the instructors themselves cited the shortage of materials for students laboratory work as one of the problems of assessment.*

# 3.2. Practices of assessing students learning outcomes …

*Mean and standard deviation of types of test items by band/field of study*

| University | Bands/fields of study | Students | | | Instructors | | |
|---|---|---|---|---|---|---|---|
| | | N | $\overline{X}$ | *s* | N | $\overline{X}$ | *s* |
| Adama | Engineering and Technology | 113 | 54.01 | 10.11 | 13 | 52.46 | 7.94 |
| Addis Ababa | Social Sciences and Humanities | 82 | 46.78 | 9.79 | 34 | 47.46 | 13.20 |
| Bahir Dar | Natural and computational Sciences | 70 | 46.46 | 10.55 | 20 | 44.98 | 6.43 |
| Hawassa | Agricultural and life Sciences | 85 | 49.83 | 8.88 | 17 | 53.06 | 8.19 |
| Jimma | Medicine and health sciences | 66 | 52.16 | 10.47 | 23 | 54.50 | 6.81 |
| Mekelle | Business and Economics | 71 | 48.66 | 6.94 | 18 | 49.97 | 9.66 |
| **Grand mean** | | **487** | **50.00** | **10.00** | **125** | **50.00** | **10.00** |

# 3.2. Practices of assessing students learning outcomes …

*Summary of ANOVA for types of test items for assessment of learning outcomes*

| Respondents | Source of variations | SS | df | MS | F | p |
|---|---|---|---|---|---|---|
| Students | Between Groups | 3974.40 | 5 | 794.88 | 8.69 | 0.00 |
| | Within Groups | 43999.52 | 481 | 91.48 | | |
| | **Total** | 47973.91 | 486 | | | |
| Instructors | Between Groups | 671.49 | 5 | 134.30 | 1.35 | 0.25 |
| | Within Groups | 11809.97 | 119 | 99.24 | | |
| | **Total** | 12481.456 | 124 | | | |

# 3.2. Practices of assessing students learning outcomes …

Use of assessment results for improving students learning

| Item | Students | | Instructors | |
|---|---|---|---|---|
| | Highest rating | % | Highest rating | % |
| Revise what has been taught | Often | 24.8 | Frequently | 25.6 |
| Provide remedial instruction | Never | 25.9 | Rarely | 30.2 |
| Provide enrichment instruction | Never | 23.5 | Rarely | 39.3 |
| Improve methods of teaching | Often | 30.2 | Frequently | 25.6 |
| Motivate students learning | Often | 25.6 | Frequently | 25.6 |
| Assign grades | Frequently | 23.7 | Frequently | 46.5 |

## Interview responses of students on use of assessment

"The assessment mechanism is less in its contribution to the quality of education. I feel that no one uses its results to take measures and improve students learning or the program. It is used to grade students and make final decision. It is not used to give feedback to the students" (GS-2).

***Comment:*** *Instructors seems to neglect the role of assessment for improving students learning.*

# Instructor responses in the use of assessment results

"I use students' results as means to help students' learning. It helped me to improve my methods of presentation. I feel that the assessment I use has made students learn better and this contributes its share to improve the quality of education." (GI-1)

***Comment:*** *It seems that is not for improving student learning but for improving instruction.*

# 3.2. Practices of assessing students learning outcomes …

*Mean and standard deviation of use of assessment results by band/field of study*

| University | Bands/fields of study | Students | | | Instructors | | |
|---|---|---|---|---|---|---|---|
| | | N | $\bar{X}$ | S | N | $\bar{X}$ | S |
| Adama | Engineering and Technology | 111 | 50.65 | 9.52 | 12 | 50.19 | 13.55 |
| Addis Ababa | Social Sciences and Humanities | 85 | 49.09 | 10.79 | 34 | 50.15 | 9.10 |
| Bahir Dar | Natural and computational Sciences | 69 | 50.22 | 11.12 | 20 | 46.78 | 8.26 |
| Hawassa | Agricultural and life Sciences | 83 | 47.82 | 8.54 | 17 | 47.69 | 11.14 |
| Jimma | Medicine and health sciences | 64 | 48.55 | 9.93 | 23 | 52.95 | 9.12 |
| Mekelle | Business and Economics | 71 | 53.60 | 9.38 | 18 | 51.57 | 10.54 |
| **Grand mean** | | **483** | **50.00** | **10.00** | **121** | **50.00** | **10.00** |

# 3.2. Practices of assessing students learning outcomes …

*Summary of ANOVA for the use of the results of assessment of learning outcomes*

| Respondents | Source of Variations | SS | df | MS | F | p |
|---|---|---|---|---|---|---|
| Students | Between Groups | 1567.89 | 5 | 313.58 | 3.22 | 0.01 |
| | Within Groups | 46524.33 | 477 | 97.54 | | |
| | **Total** | **48092.22** | **482** | | | |
| Instructors | Between Groups | 544.06 | 5 | 108.81 | 1.09 | 0.37 |
| | Within Groups | 11755.94 | 118 | 99.63 | | |
| | **Total** | **12300.00** | **123** | | | |

## 3.2. Practices of assessing students learning outcomes …

*Assessment Approaches used by instructors*

| Type of test items | Students | | | | Instructors | | | |
|---|---|---|---|---|---|---|---|---|
| | Yes | | No | | Yes | | No | |
| | N | % | N | % | N | % | N | % |
| Norm-referenced | 279 | 57.8 | 204 | 42.2 | 111 | 82.8 | 23 | 17.2 |
| Criterion-referenced | 174 | 37.0 | 296 | 63.0 | 70 | 52.2 | 64 | 47.8 |
| Formative | 96 | 20.7 | 368 | 79.3 | 25 | 18.7 | 99 | 74.4 |
| Summative | 375 | 80.8 | 89 | 19.2 | 34 | 25.6 | 109 | 81.3 |

# 3.2. Practices of assessing students learning outcomes …

*Mean and standard deviation of assessment of students learning outcomes by academic rank*

| Variables | Academic Rank | N | $\overline{X}$ | $s$ |
|---|---|---|---|---|
| Types and frequent use of Assessment stargazes | Lecturer | 37 | 53.47 | 8.31 |
| | Assistant Professor | 60 | 49.77 | 8.39 |
| | Associate Professor and above | 31 | 46.27 | 13.27 |
| | **Grand mean** | **128** | **50.00** | **10.00** |
| Types of test items | Lecturer | 36 | 53.33 | 6.54 |
| | Assistant Professor | 58 | 49.07 | 11.03 |
| | Associate Professor and above | 30 | 47.52 | 10.60 |
| | **Grand mean** | **124** | **50.00** | **10.00** |
| Use of assessment results for improving learning | Lecturer | 35 | 50.21 | 9.54 |
| | Assistant Professor | 58 | 49.90 | 9.68 |
| | Associate Professor and above | 30 | 49.88 | 11.53 |
| | **Grand mean** | **123** | **50.00** | **10.00** |

# 3.2. Practices of assessing students learning outcomes …

*Summary of ANOVA for assessment of learning outcomes by academic rank*

| Variables | Source of Variations | SS | df | MS | F | p |
|---|---|---|---|---|---|---|
| Types and frequent use of assessment stargazes | Between Groups | 879.23 | 2 | 439.62 | 4.61 | 0.01 |
| | Within Groups | 11919.57 | 125 | 95.36 | | |
| | **Total** | **12798.80** | **127** | | | |
| Types of test items | Between Groups | 634.50 | 2 | 317.25 | 3.28 | 0.04 |
| | Within Groups | 11694.01 | 121 | 96.65 | | |
| | **Total** | **12328.51** | **123** | | | |

# 3.2. Practices of assessing students learning outcomes …

*Mean and standard deviation of assessment of students learning outcomes by teaching experience*

| Variable | Teaching experience in years | N | $\overline{X}$ | $s$ |
|---|---|---|---|---|
| Types and frequent use of assessment stargazes | 10 and below | 61 | 50.65 | 8.23 |
| | 11-19 | 27 | 51.29 | 9.77 |
| | 20-29 | 27 | 47.91 | 12.52 |
| | 30 and above | 9 | 46.08 | 9.18 |
| | **Grand Mean** | **124** | **49.86** | **9.72** |
| Types of test items | 10 and below | 60 | 51.42 | 10.11 |
| | 11-19 | 26 | 50.32 | 10.08 |
| | 20-29 | 25 | 46.90 | 8.80 |
| | 30 and above | 9 | 47.28 | 10.76 |
| | **Grand Mean** | **120** | **49.93** | **9.95** |
| Use of assessment results for improving learning | 10 and below | 60 | 49.88 | 8.90 |
| | 11-19 | 26 | 50.44 | 11.04 |
| | 20-29 | 24 | 50.52 | 11.32 |
| | 30 and above | 9 | 45.46 | 10.49 |
| | **Grand Mean** | **119** | **49.80** | **9.98** |

# 3.2. Practices of assessing students learning outcomes …

*Summary of ANOVA for assessment of learning outcomes by teaching experience*

| Variables | Source of variations | SS | df | MS | F | p |
|---|---|---|---|---|---|---|
| Types and frequent use of assessment stargazes | Between Groups | 324.84 | 3 | 108.28 | 1.15 | 0.33 |
| | Within Groups | 11300.15 | 120 | 94.17 | | |
| | **Total** | **11624.99** | **123** | | | |
| Types of test items | Between Groups | 429.44 | 3 | 143.15 | 1.46 | 0.23 |
| | Within Groups | 11353.56 | 116 | 97.88 | | |
| | **Total** | **11783.00** | **119** | | | |
| Use of assessment results for improving learning | Between Groups | 193.03 | 3 | 64.34 | 0.64 | 0.59 |
| | Within Groups | 11548.92 | 115 | 100.43 | | |
| | **Total** | **11741.95** | **118** | | | |

# 4. Summary and Recommendations

**Summary**

**Purpose**

- The purpose of the study is to explore the practices and challenges of instructors' in the assessment of students learning outcomes.

- **Methods**

- Dominantly quantitative approach complemented with qualitative approach.
- The design was descriptive and the method was survey.
- Sample: 150 instructors and 600 students were sampled of which data was collected from 131 instructors and 525 students.
- Questionnaire and interview guides were used to collect data.
- Instruments were pilot-tested.
- Percentages, descriptive statistics and ANOVA were used to analyses the data.

# 4. Summary and Recommendations …

## Summary of findings

- Instructors were using variety of instruments such as individual and group projects and presentations, short and long term essays for assessing learning outcomes. However, teachers were not using computational problems, portfolios, and laboratory reports.
- Instructors in the band of business and economics use uses significantly better strategies for assessing students learning outcomes than Natural and computational Sciences and Agriculture and Life Sciences.
- Instructors tend to use frequently traditional short and long essays primarily for grading students and not for improving instructions.
- Lecturers tend to use more frequently variety assessment strategies than associate professor and professors.
- Assessment of learning is mostly used for grading students with normative approach.
- Instructors were not in use of assessment results to improve student learning.
- Most of the instructors use only mid and final examinations to assess students learning.

# Practices of assessing students learning outcomes …

## Recommendations

1. Assessment should not be used only for grading students. Instructors should emphasize the use of assessment results for improving students learning.
2. The approach for assessment of learning outcomes should be criterion referenced approach where students' knowledge, skill and attitudes are judged with reference to standard of success.
3. Emphasis should be given for the projects and portfolios in the assessment of learning outcomes.
4. There seems to be less involvement in the assessment of students learning by associate professors and professors on the as compared to lectures. Therefore, more time and attention should be given associate professors and professors as well. As it is clearly stated in the revised legislation, a unit should be established in charge of following the implementation assessment guidelines by the instructors.