

# The Stakeholders' Perception on the Reliability and Validity of Internal Assessment

Rajendra Khanal, PhD<sup>1</sup>

## Abstract

*This article presents the stakeholders' perceptions of the reliability and validity of the internal assessment system at Masters' level under the faculty of Education, Tribhuvan University. It is based on mixed-method research design as well as primary and secondary data sources. For this purpose, the purposive sampling technique was applied to select 100 master's level students studying in semester system and 25 teachers who were involved in teaching at the same level. The respondent views and perceptions have been analyzed and interpreted as descriptively and analytically. The findings depict that only 42 percent of students and 48 percent of teachers were agreed with the reliability of the score achieved from internal assessment. It can say that their level of satisfaction has signified as a weaker position. Likewise, 38 percent of students and 36 percent of teachers agreed on the tools used in internal evaluation, and their level of satisfaction is seen as very poor. However, they agreed that the tools used for internal assessment possess all sorts of validity. The findings conclude that internal assessment doesn't have a higher level of reliability and validity.*

**Keywords:** Internal assessment, semester system, final evaluation, social context, reliability and validity

## Introduction

The semester system has been practised in the education system in many developing countries like Nepal. It is considered a successful and learner-friendly educational program. The semester system's educational program has designed short-period academic sessions, lasting six months, so this program is useful for both teachers and students. The semester system has expected to help students achieve better learning attainment as the academic session changes rapidly, and they receive immediate support from the internal and external examination system. The semester system has also been important to achieve expected learning outcomes and to develop realistic qualifications for the learners. Tribhuvan University (TU) has implemented a semester system educational program at the Master level from 2015. An academic session of

1 Dr. Khanal is the lecturer of Education at Central Department of Education, TU, Kirtipur. Email: sauravabhir@gmail.com

the semester at Tribhuvan University has also completed within six months. TU has emphasized quality education by improving teaching-learning activities, examination system, and student participation through the semester system. TU has arranged internal and external examination practices in the semester system for student assessment.

Student assessment is an integral part of the semester system. It comprises both internal and external assessment systems. It helps to develop a desirable change in the capabilities of learners. Tribhuvan University has allocated 40 percent marks for internal assessment and 60 percent for external evaluation. It shows that Tribhuvan University has also given pace to both internal and external assessments to improve learning achievement at the master level throughout the semester systems. The internal assessment has been perceived as a mandatory component of the semester system. It motivates both the teachers and learners to achieve better learning outcomes.

The internal assessment has also been considered as a social context of the learning environment. It employs to enhance the effective interaction and communication among the learners and educators to attain learning objectives. It provides an opportunity for both frequent interactions among the learners and teachers to exchange their feedbacks. Fultcher and Davidson (2011) argue that internal assessment is the context of cooperative efforts in the teaching-learning process (p. 29). Its practical significance has also found to develop competencies and eliminating problems. Vitello & Williamson (2017) cited the Department for Education's definition for internal assessment as follows. "There can be seven variants of internal assessment, which differ in terms of the level of control that the awarding body has over three stages of the assessment process: task setting, task taking and task marking" (p. 543).

The internal assessment has become an integral part of the semester system, and it is a complement of the final evaluation. It helps assess students' classroom performance by evaluating learning progress and effectiveness of the teaching-learning environment. It has also directed learners to attain the curricular objectives. The internal assessment additionally motivates the learners towards learning (Khaniya, 2013, p. 111). Pathak (2012) claimed that formative assessment helps to entail learning-related difficulties in instructional units. All these arguments have explained the relevance of internal assessment to improve quality of education in the semester system.

The faculty of Education (FoE) has also implemented the same provision of the semester system designed by Tribhuvan University at the masters in Education (M.Ed.) level. The marking system of the internal assessment has also been allotting under different heads. For example, five marks for attendance, five for student participation in teaching-learning activities, and 30 for the paper-pencil test, practical test, and project work having ten for each (FoE, 2015). Other activities like writing a term paper,

reviewing a book, developing a research proposal on a particular topic, reviewing an article, taking the internal test or unit test, assessing knowledge or capacity, etc. have also been included in the curriculum.

Besides, home assignments, library work, book or article reviews, progress reports, prognostic tests, etc. have also been included in the internal assessment. However, the effectiveness, reliability, and validity of internal assessment practices have not been included either in the curriculum or academic studies. It requires innovative ideas and the participation of stakeholders to develop internal assessment tools and techniques. Their views and perceptions are also important to implement internal assessment practice in the classroom. It is the gap in this paper, and fulfilling this, this paper attempts to analyze the stakeholders' perception regarding the reliability and validity of internal assessment in the semester system. For this purpose, empirical evidence has been used to analyze the validity and reliability of the internal assessment practice in the semester system at the Master in Education (M. Ed.) level of Tribhuvan University.

## **Materials and Methods**

The paper follows a mixed-method research design with the mixing of QUAL and QUANT approaches. The analytical method has been adopted from both the teachers' and students' perceptions. It presents in both descriptive and comparative perspectives in the process of internal assessment. In this article, students studying at M Ed level and teachers teaching at the same level are called stakeholders. The study population included both the teachers and students of the constitutional campuses of Tribhuvan University within Kathmandu valley who taught the M Ed level in the semester system under the Faculty of Education. The University Campus Kirtipur, Mahendra Ratna Campus Tahachal, and Sanothimi Campus Bhaktapur have been selected for the study area. The students enrolled in the second, third, and fourth semesters in the academic year 2018/19 and the teachers who taught at the same academic year were considered the study population. Twenty-five teachers and 100 students were sampled purposively.

Primary data were collected from the sampled students and teachers. Their views and perceptions were collected from in-depth interviews, focus group discussion and key informant survey. Secondary data were collected from the official records, reviews of journals and articles, e-resources and archive documents.

The opinion of the respondent has been devised in order to accumulate the teachers' and students' perception on the concerns of internal assessment. The statements having similar nature were prepared and the perception of the respondents has been explored on those statements. The statements related to the validity and reliability were prepared at first then they were duly used to elicit teachers and students' perception.

Both descriptive and analytical methods were applied in data analysis. Individual experiences, perceptions and views have been analyzed from the Likert (1-5) scale, such as *fully agreed, agreed, undecided disagreed and fully disagreed*. The statistical data have been presented in the form of pie chart and figures where descriptive statistics has been used to present the teachers' and students' points of view.

## **Results and Discussions**

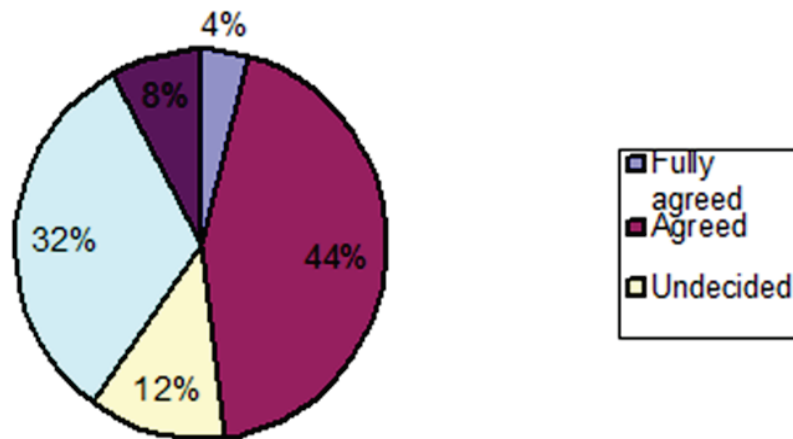
In this section, I have presented the results of my study and linked it to the theory. The discussion would be conducted on the basis of literature as well as the reflection of the researcher.

### **1. Reliability and Validity in Evaluation**

In this paper, the quality of being consistent, pure, and permanent is considered reliable. The reliable score or result of any evaluation system has legitimated in the academic field. Reliability assumes relevancy, acceptability, consistency, and trustworthiness in the tools/means used in the evaluation. The score of the tests with similar nature has reflected consistency and exact at different times. Thus, reliability can be used to measure the quality of a test item. If a test administered at a different time gives a similar result, then the quality of such a test in evaluation is reliable (Khanal & Adhikari, 2068, p. 150). Students' achievement becomes reliable if the test or assessment is reliable. The reliable achievement of the students' score also ensures the relevancy of the variables used in the field of academia, such as internal assessment practices.

Validity has also been considered as an integral part of the quality test or an assessment. The assessment means or tools have considered valid when it helps to attain desired objectives. It has focused on the practicality of constructing assessment tools. It is assumed that the representation of questions becomes valid when it would assess or evaluate students' learning outcomes through the knowledge, skill, and expected learning behavior of the students. Validity is an integral aspect of an assessment that measures the qualities as specified in objectives that which to assess while constructing the test (Khanal & Adhikari, 2068, p. 158). However, content validity is not sufficient to assess the quality of the assessment. Thus, validity is essential to assess the content, knowledge, and learning experiences specified in the curriculum for instructional evaluation. The ability to recognize, analyzes, and predicts the learners' expected outcomes and creative skills have also been included in the validity. Thus, validity has been conceived as a central theme employing to assess the quality of testing.

## 2. Reliability in Internal Assessment

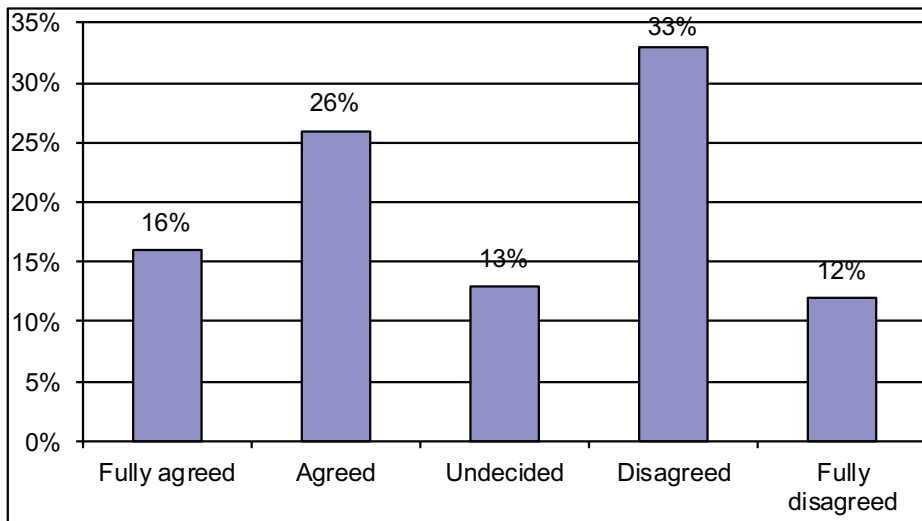


**Figure 1 : The responses toward the reliability of internal assessment**

Semester system has been taken as a modern and learner centered academic approach or program in which an academic year is segmented into two halves and the instructional and evaluative means are modified so as to help students develop autonomy and creativity. The perception on the reliability of internal assessment system adopted in semester has been taken in the present study. The teachers' responses are taken by devising a statement that is 'the score achieved from internal assessment is reliable'. The responses accumulated from the teachers' project that the score achieved by the students through internal assessment is reliable to the lesser extent. Out of the total respondents, 48% teachers are seeming to be positive whilst 52% of them are not positive on it. The responses of the teachers on the reliability of score achieved can be presented in Figure 1.

The teachers' responses related to reliability are presented in Figure 1. The figure indicates that 41 percent of teachers have agreed whereas 44 percent have expressed their agreement with this statement. Similarly, 12% of teachers found undecided and 32 percent in disagreement. Even 8% of them fully disagreed. This sort of score proves that the level of reliability is weaker or not satisfactory one. Though the teachers themselves give marks to the students, they are not agreed that the marks from internal assessment have reliability. It proves that the teachers do not believe that the marks they provided to the students by assessing their capacity is reliable. It has raised a serious question on the marks from internal assessment. Moreover, it implies the essence of identifying the reasons or affective factors behind the existence of such unbelievable/questionable condition of internal assessment's marks provided by the teachers themselves.

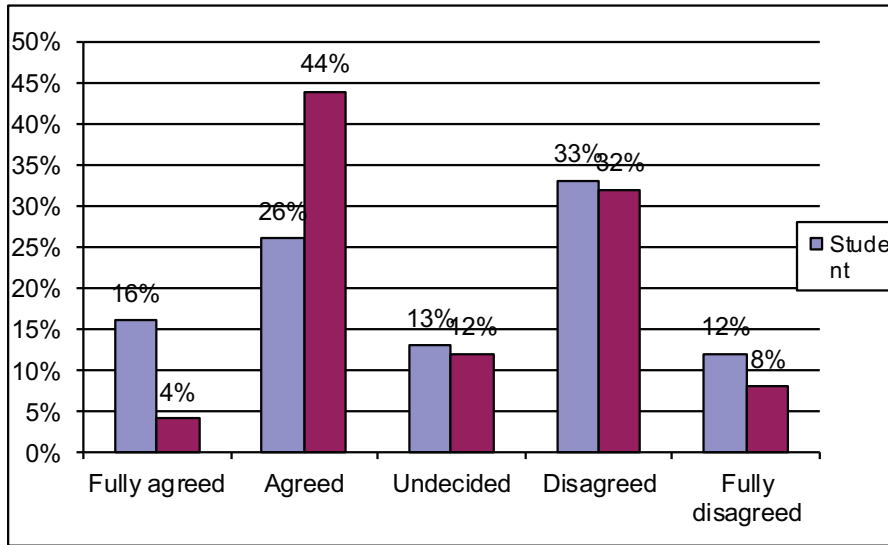
In the same line, the data also shows that even the students are not satisfied to the marks they achieved from internal assessment. In total, only 42% of them are positive whilst 58% of them disagreed that the marks they achieved is reliable. The responses on the reliability of achieved marks can be presented in Figure 2.



**Figure 2 : Responses on the reliability of internal assessment**

The figure presents the opinion of the students towards the reliability of internal assessment. Accordingly, 16% of the students in total showed their absolute agreement and 26 of them showed their agreement to the statement. In this way, 42% of them opined positively in total. This data proves that the level of reliability is not good or is questionable. Similarly, 13% students in total are undecided whilst 33% of them expressed their disagreement to the reliability of marks. Even 12% students are fully disagreed to the statement. So, it has raised a severe question on the marks achieved from internal assessment. It further implies that there is the need of exploring the affective factors that lead to the state of distrust in the marks achieved by the students in order to make internal assessment reliable.

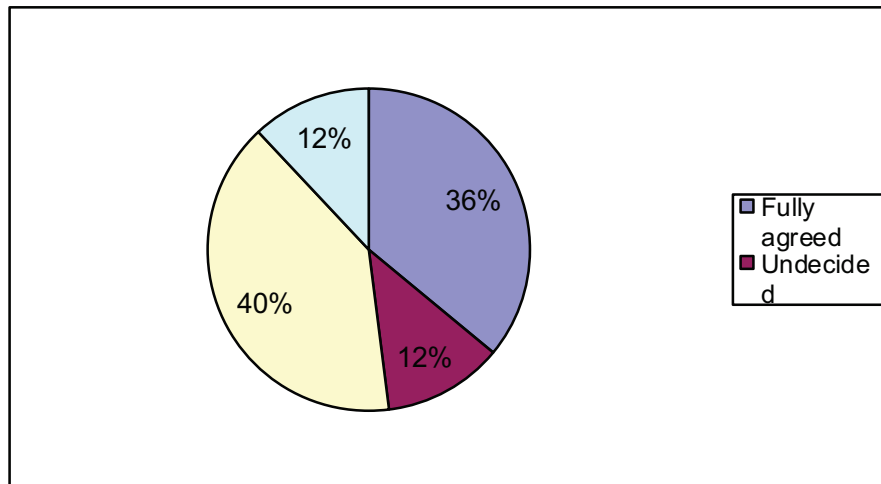
The marks achieved by the students from internal assessment have been found relatively less reliable. In aggregate, comparing, the responses, only 48% teachers and 42% students seemed to be positive whilst others disagreed the statement. The responses on the reliability of the marks achieved have been presented in Figure 3.



**Figure 3 : Perception towards the Reliability of Internal Assessment**

The figure depicts the responses of both the teachers and students towards the reliability of internal assessment. Only 4% teachers and 16% students are fully agreed and 44% teachers and 16% students showed their agreement to the statement, 'internal assessment is reliable'. In contrast, 12% teachers and 13% students in total remained uncertain, whereas 32% teachers and 33% students presented their disagreement to the reliability of the marks. Moreover, 8% teachers and 12% students fully disagreed the statement. This shows the similar nature of perception of both the teachers and students. Amongst the students, 60% of them expressed less positive responses than those of the teachers. Though the disagreement of the students is conceivable, the teachers who assessed the students and provided the marks do not believe that the mark is reliable. This is really unpredictable concern. Teachers' such perception has raised a serious question on the marks achieved or provided in internal assessment. Neither the students are satisfied nor the teachers considered the marks as being reliable. This proves that internal assessment does not have reliability. So, the result implies that, there is an intense need of identifying the situation like this or even exploring the affected sources.

### 3. Validity in Internal Assessment



**Figure 4 : The responses towards the validity of internal assessment**

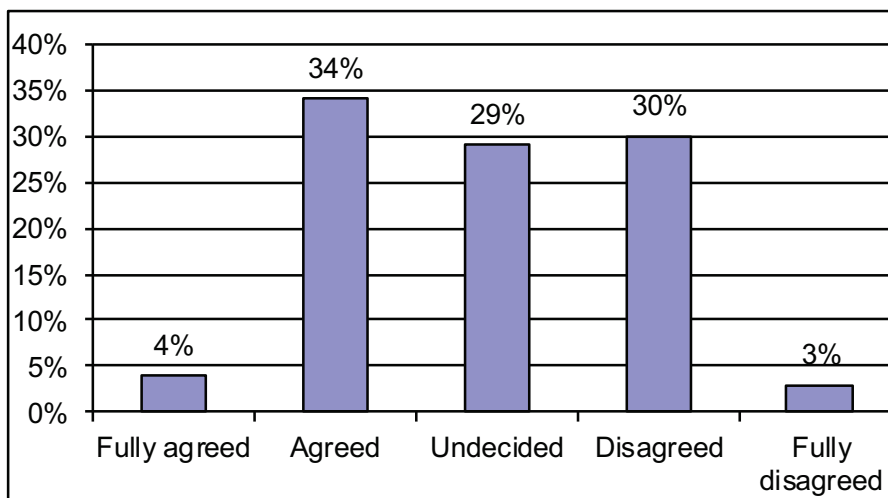
All forms of validities are also expected in the internal assessment implemented at masters' level in semester system under Tribhuvan University. In the present study, the perception of the teachers and the students has been elicited by developing a statement related to the validity of assessment tools. The tools of internal assessment system include: project work, presentation, internal tests, etc. Thus, the opinion has been elicited focusing up on those tools. Perception of the teachers so far obtained clarified that validity of the tools used in the process of internal assessment system is uncertain. Analysis of the opinions in this indicator entails very weaker state of validity. It is because only 36% of teachers are positive and 64% of them in total are negative to the statement which proves weaker image of validity in Figure 4.

The teachers' opinion regarding the validity has been presented in the above chart. As per the data, 36% teachers in total accepted that there is validity in the tools of assessment. In contrast, 12% teachers are undecided and equal numbers of them are fully disagreed to the statement developed for validity whereas 40% of them disagreed. Since the majority of teachers, i.e. 52% in total disagreed and 12% of them remain uncertain, there is the problem of validity in the tools used in assessment. In this way, 64% teachers in total do not accept that there is validity in the tools used in internal assessment. Such a figure clarifies that there is a big question in the validity of internal assessment system.

Since the teachers presented themselves in the poll that the tools of internal assessment are beyond the notion of validity, it is important to have an intensive discussion among the stakeholders regarding the development and the implementation of the tools for the process of internal assessment. Similarly, there is the need of exploring the reasons



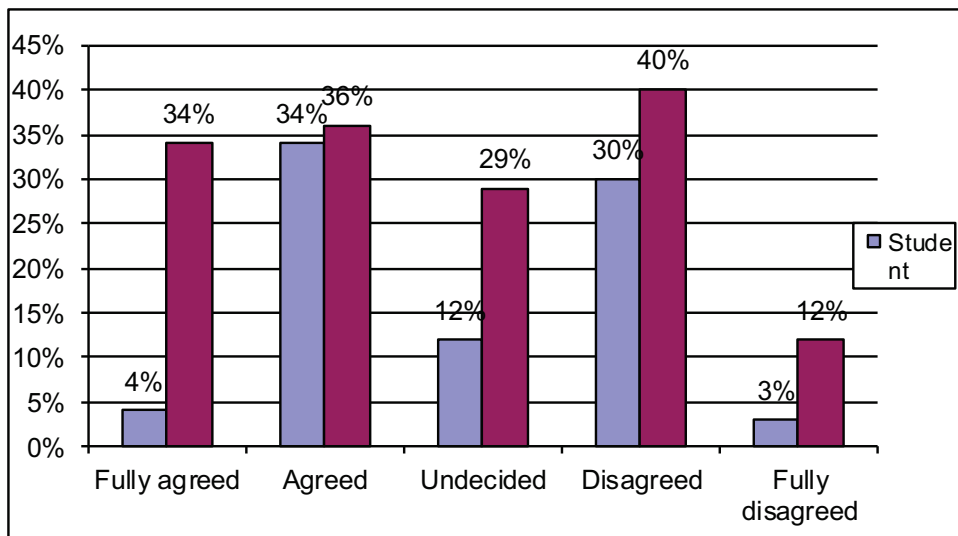
being invalid and to well inform the teachers on that very concern.



**Figure 5: The opinion towards the validity in internal assessment**

The students' responses obtained in the statement also clarify the uncertainty regarding the validity of the process of internal assessment. Analyzing the responses so far obtained in this statement, we come to see the vulnerable state of validity in assessment. The data representing 38% in total in positive and 62% in negative sense also portrays the worst picture of validity in Figure 5. The students' opinions related to the validity of internal assessment have been presented in the above figure. According to the data in the figure, 38% students in total have accepted that the assessment has validity, 29% of them have not given decisive opinion and 3% of the students are fully disagreed to the notion of validity. Since the majority, i.e. 62% of the students in total do not accept that there is validity in the tools used in internal assessment or are not positive towards the statement; a big question has been raised to the validity of internal assessment. The data implies the need of intensive discussion up on the process of internal evaluation so as to make the assessment more valid.

The comparative analysis of the responses obtained to the notion of validity projects very worse state. The data representing 36% teachers and 36% students in total in the line of validity of the tools and the rest of 64% teachers and 62% students in total against it proves the vulnerable state of validity in Figure 6.



**Figure 6 : The responses towards the validity in internal assessment**

Both the teachers and students' perception regarding the validity of assessment has been presented in the figure above. Accordingly, 4% students in total are fully agreed whilst none of the teachers showed their fuller agreement to the statement that there is validity in internal assessment. Likewise, 34% students and 36% teachers in total have shown their agreement on the tools of assessment as being valid. In this way, 38% students and 36% teachers are positive to the statement. In contrast, 12% teachers and 29% students deny to give decisive opinion whilst 40% teachers and 30% students in total disagreed on it. And 12% teachers and 3% students seem to fully disagree towards the statement developed on the concern of validity. The figure of 52% teachers and 33% students in total being disagreed to the statement, and 12% teachers and 29% students being agreed signifies that there is a severe problem in the validity of internal assessment.

In this way, a big question has been raised in the validity of internal assessment. It is because the majority of teachers and students, i.e. 64 and 62% respectively are unable to accept the statement that there is validity in the tools of internal assessment. It is quite surprising that the teachers themselves directly develop and use the tools but consider them as being not valid much. This is a highly sensible matter. Thus, it is almost essential to have in-depth discussion upon the validity of tools to be used among the stakeholders at first and then to revise/reframe the tools of assessment.

The results of the study enable the researcher to conclude that the notions of validity and reliability in internal assessment system are questionable. The figure of 42% students and 48% teachers in total being positive to the statement 'the score obtained

from internal assessment is reliable' depicts the poor satisfaction of the teachers and students, both. In the same way, the figure of 38% students and 36% teachers in total being positive to the statement that 'the tools used in internal assessment system possess all forms of validity' signifies their poor level of satisfaction or acceptance. Though the teachers themselves are directly involved in devising the tools and in providing marks to the students, they themselves express their doubt on the notion of validity and reliability. It has explicitly clarified the professional insensibility and procedural complexity inherent within internal assessment.

### **Conclusion**

The results of the study enable the researcher to conclude that the notions of validity and reliability in internal assessment system are questionable. The figure of 42% students and 48% teachers in total being positive to the statement 'the score obtained from internal assessment is reliable' depicts the poor satisfaction of the teachers and students, both. In the same way, the figure of 38% students and 36% teachers in total being positive to the statement that 'the tools used in internal assessment system possess all forms of validity' signifies their poor level of satisfaction or acceptance. Though the teachers themselves are directly involved in devising the tools and in providing marks to the students, they themselves express their doubt on the notion of validity and reliability. It has explicitly clarified the professional insensibility and procedural complexity inherent within internal assessment.

### **Recommendations/Implications**

Similar kind of perception of the teachers and students has been derived regarding the reliability and validity of internal assessment from this study. The teachers and students both portrayed very weak state of reliability and validity in their opinions though they are taken as the integral qualities of a good assessment. A few constructive/reformative suggestions have been presented from the obtained responses in the study. The weaker image of the reliability of the marks obtained from internal evaluation has been seen. The teachers must be more responsible and conscious enough towards the issue of reliability and validity as they themselves are involved in the construction of tools and scoring the learners' achievement, despite their uncertainty in the score and the process of internal assessment. Undoubtedly, the teachers must be sensible in case of any administrative huddles or the influence in case of the reliability of the score obtained by the students. It is the teacher who has to control or lesion the negative impulses to create any unwanted obstruction. Both the teachers' and students' perception has been seen quite disappointing to the notion of validity of the tools used in internal assessment. That's why, the teachers must also be thoughtful in order to maintain the validity of the tools to be used in the internal assessment.

## References

- Ali, J.G. (1996). *Webster dictionary*. USA: Otenheimer Publishers.
- Chongbang, K.B. (2014). *Comparative study of semester system and annual system of Faculty of Education*. T.U. A research report submitted to the office of Dean. Faculty of Education.
- Faculty of Education (2072). *Curriculum of Nepali education, Masters level*, Faculty of Education, T.U., Kirtipur.
- Fultcher, G. & Davidson, F. (2011). *Language testing and assessment*. London: Routledge.
- Heaton, J. B. (1975). *Writing English language tests*. London: Longman.
- Khanal, P. & Adhikari, N.P. (2068). *Curriculum and evaluation*. Kathmandu: Sunlight Publication.
- Khaniya, T. R. (2013). *Examination for enhanced learning*. Educational and Developmental Service Center.
- Lado, R. (1965). *Language testing*. London: Longmans.
- Lamsal, R. & Khanal, R. (2075). *Research methodology*. Kirtipur: Sunlight Publication.
- Majumdar, M. (2010). *Introduction of semester system in Indian College*. Asam University: A seminar Paper.
- Niure, D. P. (2070). *Curriculum and evaluation*. Kathmandu: Quest Publication.
- Pathak, R. P. (2012). *Measurement and Evaluation in Education*. Delhi: Pearson.
- Poudel, L. (2074). *Semester system: Problem and Measures to improve*. A concept paper presented in an interaction program in Kirtipur organized by Nepal progressive professors Association.
- Renalds, C. R. & Others (2018). *Measurement and assessment in education*. India: Pearson.
- Sharma, K. P. & Poudel, M.P. (2068). *Contexts of Nepali language Teaching*. Kathmandu: Students Book Store.
- Tripathi, S. & Others (2018). *T.U. Bulletin Special*. Kathmandu: Tribhuvan University.
- Vitello, S. & Williamson, J. (2017). Internal versus external assessment in vocational qualifications: A commentary on the government's reforms in England. *London Review of Education, Volume 15 (3)*, pp. 536-548.
- Weir, C. J. (1996). *Communicative language testing*. London: Blackwell Publication.
- Yusuf, A. & Hasim, M. (2012). A case study of annual and semester system of examination in government college of management science *International Journal of Academic Research in Business and Social Science*. Vol. 2 (53-73).