Development and Validation of a Longitudinal Assessment Model using Normalised Change to Improve the Quality of Educational Outcome Standards

Fasasi, R.O*

Department of Information Technology, Faculty of Accounting and Informatics,
Durban University of Technology
fasasi@outlook.com

Heukelman D

Department of Information Technology, Faculty of Accounting and Informatics,
Durban University of Technology
deleneh@dut.ac.za

Doi:10.5901/mjss.2014.v5n27p564

Abstract

The concept of "knowledge assessment" is increasingly utilized, but not yet well defined. Despite its vagueness, or variability according to the people who utilize it, the concept of knowledge assessment has been imposed in the management literature in general and in the inculcated literature in particular over the past thirty years. However, professional competence remains elusive and equivocal because of its fragility resulting from the distinction between its importance and its low calibre of definition. The longitudinal assessment model is therefore developed to clearly define the usefulness of assessment, using normalised change to improve the quality of education standards. Though educational assessment provides information on whether teaching/knowledge has been prosperous, the information it provides has a number of potential different audiences, whose precise requisites may vary. In education, it is essential for instructors to gather information on how pupils', understanding, knowledge and skills are developing, both to appraise how they should adjust their teaching method and to determine what kind of feedback is needed to upgrade pupils. Furthermore, education administrators as well as policy makers need to be provided with additional broader information on the quality of education on a continent or in a country. Government also requires information which can help understand their workers' achievements and constraints. Pupils themselves require knowing how they are progressing and how to improve their performance, and to be protected from the potentially demotivating effects of negative assessment. The aim therefore of the Longitudinal Assessment Model using Normalised Change is therefore to provide a wider spectrum of reporting, which could offer more insight into the actual performance of the different parties.

Keywords: *progress assessment, longitudinal study, normalised gain, normalised change, Pre/Post-Test assessment*

1. Introduction

Assessing knowledge in higher education poses challenges. There seems to be various suggestions on the best ways of conducting an assessment according to Suskie (2010) and Hake (2007). A logical approach often adopted is Pre-assessing and Post-assessing as described by Bond (2009). Knowing what skill or knowledge an individual has acquired over a giving session (term) would require an assessment of what such an individual knew at the beginning of the session and an assessment at the end of the term, using the same or similar instrument. Rogosa (1995), Willett (1997) and Willett (1989), give a theoretical basis for the technique and also proved that more sessions of both Pre-assessment and Post-assessment would significantly improve the reliability of the method.

This longitudinal assessment scheme involves breaking or dividing the main course into several distinct teaching sessions. The distinct sessions are separated by carrying out Pre-Post-teaching assessments using the Post-assessment of one session as the Pre-assessment of the next successive session. Both Pre- and Post-assessments give relevant information on a distinct area of the course's effectiveness. An overall review will be required to ascertain if exposing students to coursework materials over many sessions of teachings would increase the possibility of such students harvesting and retaining the basic knowledge. Since student performance is expected to vary across each teaching session, there is the need to continuously balance both the retention and accession efforts of students by adopting a marginal review to determine the extent of flexibility of the course design. The approach used to weigh variation in

student performance is an explanatory characteristic of any Pre- or Post-assessment strategy. The most adopted index in evaluating the variation in any group performance between the Pre-teaching assessment and the Post-teaching assessment is

$$q = \frac{\text{average score}_{\text{post-teaching assessment}} - \text{average score}_{\text{pre-teaching assessment}}}{100 - \text{average score}_{\text{pre-teaching assessment}}} \quad 1$$

Where, q is the normalised change, which indicates the variance between the mean of the assessment scores, representing the ratio of the maximum likely variance flanked by these scores. Hovland *et al.* (1949) adopted q as a ratio in measuring the effect of using Instructional films whiling teaching. Hake (1998) also adopted *q* to evaluate the effect of implementing several teaching techniques in introductory physics courses. Subsequently, Meltzer (2002) adopted q to investigate the relationship that exists between concept knowledge in physics and mathematics preparation. In these studies, it was identified that the subject having the biggest value of q has the highest rate of comprehension compared to the other subjects. Though this assessment rule could result in counterintuitive conclusions; q can however be decomposed using an alternative assessment rule:

$$q = Q - \lambda T \quad 2$$

Where Q is representing a normalised gain, quantifying the probability that an error in the groups Pre-teaching assessment is addressed and corrected on the Post-teaching assessment. T represents the normalised loss, quantifying the probability that in the group Pre-teaching assessment all correct responses are reported as incorrect in the Post-teaching assessment. λ is a non-negative parameter representing the renormalisation factor that depends on the Pre-teaching performance of the whole population.

Thus, $q = Q - \lambda T$ indicates q (the re-normalised change) as the difference between the normalised gain and the re-normalised loss, which are both non-negative parameters.

Equation (2) was derived based on the subsequent observation:

Where Q = *the number of correctly answered questions of the Post-teaching assessment and the incorrectly answered of the Pre-teaching assessment over the number of correctly answered in the Pre-teaching assessment over the number of incorrectly answered in the Pre-teaching assessment.*

T = *the number of incorrectly answered questions of the Post-teaching assessment and the correctly answered of the Pre-teaching assessment over the number of correctly answered in the Pre-teaching assessment*.

Here, the numerator Q is representing the number of questions in which pupils reflect a gain in knowledge, while the denominator is representing the maximum likelihood gain. However, the ratio Q represents the normalised gain proving the likelihood available mistakes in the pre-teaching assessment test had been addressed in the post- teaching assessment.

Furthermore, the numerator in T represents the number of questions which indicate pupils' loss of knowledge while the denominator represents the maximum likelihood loss. However, the ratio T represents the normalised loss proving the likelihood that the correct answers from the pre-teaching assessment is treated as incorrect from the post-teaching assessment. Therefore, q decomposed in Q provides an alternative assessment rule devoid of the counterintuitive conclusions as earlier stated and translates to the fact that:

The subject having the biggest value of Q (normalized gain) with the smallest value of λT (the re-normalized loss) is the more effective subject.

2. Components of Normalised Change

Normalised change for a group of N pupils who answered the assessment question M, can be stated in the following form:

$$q = \frac{\theta_{post-test} - \theta_{pre-test}}{1 - \theta_{pre-test}} \quad 3$$

Where $\theta_{post-test}$ represents the assessment score of skills acquired after the teaching session and $\theta_{pre-test}$ represents the assessment score of skills the pupils had before the teaching session. The scaling factor λ is a non-negative parameter representing the ratio of people involved in the Pre-teaching assessment process.

$$\lambda = \frac{\theta_{pre-test}}{1 - \theta_{pre-test}}, where \ \theta_{pre-test} < 1 \quad 4$$

The change in the assessment score for the skill acquired during the teaching assessment process can be

calculated by substituting equation (2), equation (3) and equation (4). This can be expressed in equation 5.

$$\theta_{post-test} - \theta_{pre-test} = Q(1-\theta_{pre-test}) - T\theta_{pre-test} \quad 5$$

3. Measurement of Course Effectiveness

The following steps are used to assess the relative effectiveness of the training programme and also to measure the progress of the subject. Figure 1 shows the set of pupils A involved in the training programme.

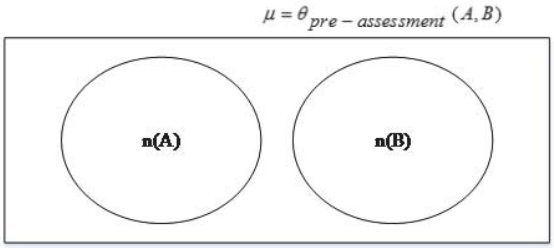

Figure 1: The Venn diagram of Control group and Experimental Group.

A Represents the control group of Pre-teaching and this is the group of the pupils who have additional knowledge in the Pre-teaching assessment. Therefore $n(A)$ is the number of pupils in the control group. B Represents the experimental group; this is the group of pupils who needed more attention on the teaching session and who is the focus of the instructor, $n(B)$ is the number of pupils in the experimental group. $n(A \cup B)$ Represents the total number of pupils involved in the assessment exercise. This is expressed in equation 6.

$$\theta_{pre}(A,B) = n(A \cup B) = n(A) + n(B) \quad 6$$

The assessment exercise is categorized into two, namely: Pre-assessment ($\theta_{pre}(A,B)$) and Post-assessment ($\theta_{post}(A,B)$). Equations 7 and 8 show the relationship between the result of Pre-assessment and Post-assessment exercise for both the control group and experimental group.

$$\lambda_{pre}(A,B) = \lambda_{A_{pre}} + \lambda_{B_{pre}} \quad 7$$

$$\lambda_{post}(A,B) = \lambda_{A_{post}} + \lambda_{B_{post}} \quad 8$$

Where $\lambda_{A_{pre}}$ represents the result of the pupils in the control group before teaching $\theta_{pre}(A,B)$; $\lambda_{B_{pre}}$ is the result of the pupils in the experimental group before teaching $\theta_{pre}(A,B)$; and $\lambda_{pre}(A,B)$ is the result of all the pupils before teaching $\theta_{pre}(A,B)$; $\lambda_{A_{post}}$ represents the result of the pupils in the control group after teaching $\theta_{post}(A,B)$; $\lambda_{B_{post}}$ is the result of the pupils in the experimental group after $\theta_{post}(A,B)$; and $\lambda_{post}(A,B)$ is the result of all the pupils after $\theta_{post}(A,B)$.

$$\delta(\lambda_{B_{post}}, \lambda_{B_{pre}}) = \lambda_{B_{post}} - \lambda_{B_{pre}} \; iff \; \lambda_{B_{post}} > \lambda_{B_{pre}} \quad 9$$

Equation 9 measures the knowledge and skills gain for the experimental group B. Equation 9 shows the measurement of the result of the Post-assessment for an experimental group with the result of the Pre-assessment of the experimental group. This is to ascertain the level of the knowledge acquired by the experimental group after the end of the training programme.

$$\delta(\lambda_{B_{post}}, \lambda_{A_{pre}}) = \lambda_{B_{post}} - \lambda_{A_{pre}} \; iff \; \lambda_{B_{post}} > \lambda_{A_{pre}} \quad 10$$

Equation 10 compares the knowledge and skills gained by the experimental group B to the knowledge and skills which the control group A has prior to and teaching.

Equation 10 shows the measurement of the result of the Post-assessment of the experimental group with the result of the Pre-assessment of the control group. This is to assess if the experimental group and the control group are at the same knowledge level. A positive value from the equation 10 will indicate that the training programme was effective and met its Predefined objectives. It is essential to measure the knowledge acquired for each group at the end of assessment exercises.

4. Longitudinal Assessment

A lot of Pre- or Post-assessment progress measurement use only one instrument session, paused or separated by similar or very similar Pre-and Post-teaching assessments, according to some research works such as McConnell *et al.* (2006), Libarkin *et al.* (2006), and Meltzer (2002). However, these single-stage approaches that depend on two assessments are associated with weaknesses as a result of their inability to harvest sufficient data capable of detecting unavoidable variations in knowledge obtained from non-perfect accession and retaining of study materials. Introducing a Pre-Test or Post-Test exercise in a single session would not be able to detect performance differences between an individual who learned a key skill and had forgotten it and an individual who never learned the same skill at all. Another weakness with the single-stage approach is its inability to distinguish between an individual who maintains Pre-teaching knowledge during the course of the session and another individual who forfeits knowledge and then relearns the knowledge again during the knowledge section or term.

Longitudinal assessment schemes monitor variations in knowledge and also fine-tune the process of assessment by integrating various single-stage schemes. The 2-point approach as shown in figure 2 is capable of detecting a one-time loss and subsequent gains or reaccessions of study material and also a formal accession and then loss of the course material. It is significant to observe that such inter-session analytic assessment (T_1) produces the Post-teaching assessment of the first stage and also the Pre-teaching assessment of the second stage.

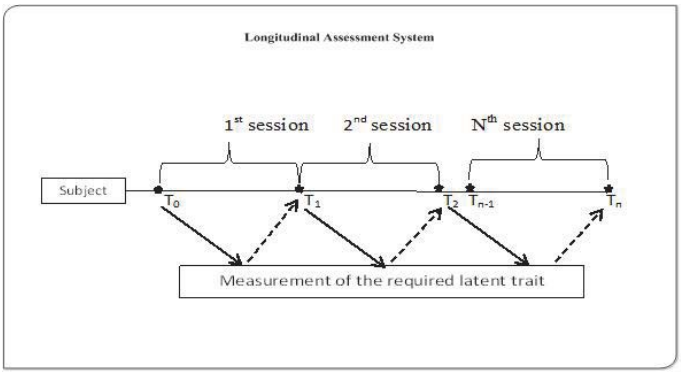

Figure 2: The System diagram of Longitudinal Assessment

Figure 2 shows the first session of a longitudinal assessment system grouped by Pre- and Post-teaching assessments T_0 and T_1. The second stage is grouped by T_1 and T_2. The analytic assessment is using relative tools developed to assess the key concepts and skills acquired.

5. Assessing The Marginal in Longitudinal System

The normalised change component is used in a marginal analysis to formulate fluctuations among performance subject to the Pre-teaching stage of every period in the longitudinal approach. Fluctuations in effectiveness between one teaching session and the next for a given course can be analysed using this method.

It is essential to note that the standard of determining effectiveness varies between sessions for marginal analysis, figure 2 showing 2-point approaches could have marginal analyses using knowledge gained and knowledge lost between T_0 and T_1 and also between T_1 and T_2 to analyse fluctuations in effectiveness for a single course. However, circumstances to improve effectiveness of the first teaching session for the second teaching session indicate that there is greater course effectiveness in enhancing knowledge for the T_1 performance than if it was compared to the T_0 performance.

Marginal analysis of a 2-point approach could also be used in comparing the effective variance of two different courses to improve knowledge related to T_0 and also to T_1. In such an instance, it is possible that either of these courses is more effective to improve knowledge related to T_0, while another is more effective to promote knowledge related to T_1.

6. Collective Analysis of Longitudinal System

The overall successive stage is what cumulative analysis formulates with their change in performance by evaluation starting from the Preliminary Pre-teaching assessment to the succeeding Post-teaching assessments, both the gains and losses. It is distinct from the marginal analysis, in that the performance on the initial T_0 is stable and based upon the measure of change along successive sessions following T_0 to T_1, T_0 to T_2, ……..T_{n-1} to T_n.

This method can be adopted to evaluate the effectiveness of certain courses along the single session between T_0 to T_1 and also how effective it is along two sessions between T_0 to T_2. With this, the students can correct their weaknesses as indicated from their performance in the basic diagnostic Pre-teaching assessment, T_0. However, the method can as well be adopted to review the effective variance of two distinct courses, to improve knowledge along the 1st two to three, and even more teaching sessions subsequent to the basic diagnostic Pre-teaching assessment T_0.

7. Sample

The illustration here adopts a longitudinal assessment model pre- and post-assessment to measure the effectiveness of a training program established by the National e-Skills Plan of Action (Olugbara et.al, 2014) in South Africa on training community development workers on ICT e-skills. The training lasted for 4 days per group in 2013 with 8 groups. The training was delivered through Blackboard Learning Management System facilitated by an instructor. The subject sample had a total range 80.2% from rural areas with 16.1% from semi-rural areas and 3.7% form urban areas. 57.8% had grade 12 educational qualifications, 41.8% had post-grade 12 qualifications and fewer than 0.6% had a NQF level 6 certificate., 58.6% of the respondents has had formal training on how to use a computer and 41.4% had not been introduced to computers in any way, 40.9% indicated to have a low competence of using computers and 48.4% indicated to be average while only 4.9% indicated to be competent and 5.8% had never used a computer before. 87.5% reportedly had access to the internet and 12.5% have no access to the internet. Over 50% of the community development workers are required to search for information daily through internet while approximately 67.5 of them indicated that their job prescription required them to telephone daily. Over 55% of them are involved in providing information to individual community members daily, while only 16% are not required to use email at their workplace and the remainder do use email to carry out their duties.

Table 1: Profile of Sample

Education Level	NQF 4: 57.8%	NQF 5: 41.8%	NQF 6: 0.6%	
Formal ICT Training	Yes: 58.6%	No: 41.4%		
Competence in ICT	High: 4.9%	Average: 48.4%	Low: 40.9%	None: 5.8%
Internet Access	Yes: 58.6%	No: 41.4%		

8. Methodology

A self-reporting assessment was conducted to evaluate the e-skills needs of CDWs' in supporting community development projects. The important factors were to determine how and to what extent ICT is currently used to support government projects for community development. The results were to be used to suggest a prototype system to be designed, based on actual pedagogy, for the community development workers to improve their ICT skills according to their needs. KwaZulu-Natal province of South Africa was chosen as a place for conducting the pilot study. KwaZulu-Natal is one of the most sparsely populated provinces of South Africa and, thus is a good representative of South African rural areas. The study population were community development workers to participate in the e-skilling project in KZN. The cluster sampling method was applied to KZN catchment in the province, which have been involved in the community development project. The self-reporting assessment was divided into two parts. In the first part, each respondent was asked to indicate the frequency of the tasks: daily, weekly and monthly, as well as whether they currently use ICT to complete the tasks. In the second part, each participant was asked to indicate their personal level of skills in the use of ICT applications by selecting from the options from 1 (No skill), 2(Limited skill) 3(Average skill) 4(Good skill) and 5 (Expert skill). The data obtained was analysed with IATA software using the partial credit model.

9. Findings and Discussion

Table 2: Frequency of tasks performed by CDWs which could be enhanced using ICT

Task	Frequency			
	Daily	Weekly	Monthly	Never
Write report	12	41	249	25
Email people	88	117	71	51
Provide information to individuals	181	64	36	46
Talk to group of people	100	134	45	48
Attend meetings	41	205	43	38
Phone people	221	43	23	40
Search for information	169	79	33	46
Provide forms for people to complete	43	97	117	70
Sms group of people	115	108	54	50
Submit a budget of financial statement	25	23	97	182

9.1 Pre-Assessment Analysis Result for Self-Assessment of Community Development Workers on ICT Skills

Table 3 shows the Pre-Test Analysis Result for Self-Assessment of Community Development Workers on ICT Skills. Item 11 in the assessment in Table 3 "To what extent can you use social media (sms, email, twitter) to facilitate mobile interaction between community members and government representatives?" produced the highest result in the EXPERT SKILLED, GOOD SKILLED and AVERAGELY SKILLED category with value of 2.4%, 16.2% and 24.2% respectively indicating that the majority of the participants in the Pre-Test Assessment has limited or no skill in the rest of the items. For item 9 "To what extent can you use a search engine (search engine in a government portal/website) to find funding opportunities, donors and development agencies?" the highest result was obtained in the LIMITED SKILLED category with a value of 29.2% indicating that 95 participants has limited skills in the area of using search engines to find opportunities, donors and development agencies. Item 17, "To what extent can you use an electronic spreadsheet to set up a financial budget" produced the highest result in the NO SKILL category, with a value of 78%, indicating that most of the participants have no knowledge or skills in using electronic spreadsheets to set up a financial budget.

Table 3: Pre-Assessment Response Frequency for Self-Assessment of Community Development Workers on ICT Skill

Item Description		Valid					
		No Skill	Limited Skill	Average Skill	Good Skill	Expert Skill	Total
ITEM 2 (Word Processor)	Frequency	156	53	90	26	2	327
	Percent	47.7	16.2	27.5	8.0	.6	100.0
ITEM 3 Use Template	Frequency	97	83	98	49	0	327
	Percent	29.7	25.4	30.0	15.0	0.0	100.0
ITEM 4 (Use power Point)	Frequency	199	59	38	31	0	327
	Percent	60.9	18.0	11.6	9.5	0.0	100.0
ITEM 5 (Presentation Software)	Frequency	216	47	53	11	0	327
	Percent	66.1	14.4	16.2	3.4	0.0	100.0
ITEM 6 (Search Engine For Crop Management)	Frequency	172	66	47	39	3	327
	Percent	52.6	20.2	14.4	11.9	.9	100.0
ITEM 7 (Search Engine For Crop Marked)	Frequency	210	60	37	19	1	327
	Percent	64.2	18.3	11.3	5.8	.3	100.0
ITEM 8 (Search Engine For Supply Chain)	Frequency	231	64	19	12	1	327
	Percent	70.6	19.6	5.8	3.7	.3	100.0
ITEM 9 (Search Engine For Dev. Agencies)	Frequency	168	95	38	24	2	327
	Percent	51.4	**29.1**	11.6	7.3	.6	100.0
ITEM 10 (Electronic spreadsheet)	Frequency	203	73	35	14	2	327
	Percent	62.1	22.3	10.7	4.3	.6	100.0
ITEM 11 (Use Social Medial)	Frequency	99	88	79	53	8	327
	Percent	30.3	26.9	**24.2**	**16.2**	**2.4**	100.0
ITEM 12 (Downloading From Website)	Frequency	144	71	73	34	5	327
	Percent	44.0	21.7	22.3	10.4	1.5	100.0
ITEM 13 (Use Really Simple Syndication)	Frequency	237	53	25	10	2	327
	Percent	72.5	16.2	7.6	3.1	.6	100.0
ITEM 14 (Use Electronic Media)	Frequency	222	55	31	18	1	327
	Percent	67.9	16.8	9.5	5.5	.3	100.0
ITEM 15 (Electronic Media To Setup Meeting)	Frequency	221	44	48	12	2	327
	Percent	67.6	13.5	14.7	3.7	.6	100.0
ITEM 16 (Setup Group Email Account)	Frequency	222	54	41	10	0	327
	Percent	67.9	16.5	12.5	3.1	0.0	100.0
ITEM 17 (Use Spreadsheet For Financial Budget)	Frequency	255	46	21	5	0	327
	Percent	**78**	14.1	6.4	1.5	0.0	100.0
ITEM 18 (Capturing of Data)	Frequency	213	63	39	12	0	327
	Percent	65.1	19.3	11.9	3.7	0.0	100.0
ITEM 19 (Creating new Networking)	Frequency	219	63	26	18	1	327
	Percent	67.0	19.3	8.0	5.5	.3	100.0
ITEM 20 (Access Government Information)	Frequency	155	67	62	36	7	327
	Percent	47.4	20.5	19.0	11.0	2.1	100.0

9.1.1 Pre-Test Scale Review Percent Score

In percent score, each bar represents a range of scores, less than 20% of the examinee scored between the percent score range of 50% and the height of each bar represents the proportion of examinees with scores at each range. In the chart in figure 3, column height shows frequency - the number of examinees in overall item score grouping. For example, over 80% of CDW received a test score between the ranges score of 0 to 5%.

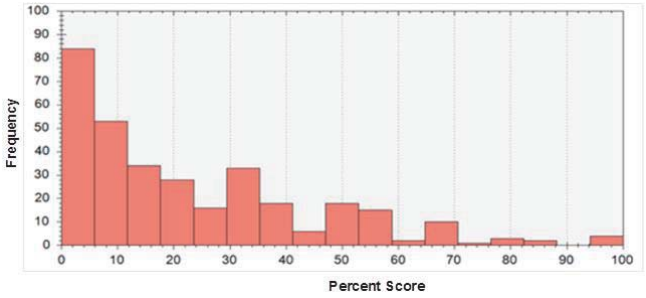

Figure 3: The Percent score for the Pre-Test assessment of Community Development Workers on ICT.

9.1.2 Pre -Test Performance Standards

In a statistical analysis, the probability value in figure 4 indicates the probability of obtaining same value for a model formulated between two hypotheses, from these hypotheses one is represented as "neutral" (or "null") while the other is placed under the hypothesis measuring. Threshold is set previously and the probability is less than the default threshold (traditionally 5% or 1%). However, we can accept the measure hypothesis as valid and reject the neutral hypothesis. Additionally, the performance standard result for this assessment is valid as the threshold is 1.43, with the mean of 1.14, and the standard deviation is 0.79. However, the probability is less than threshold.

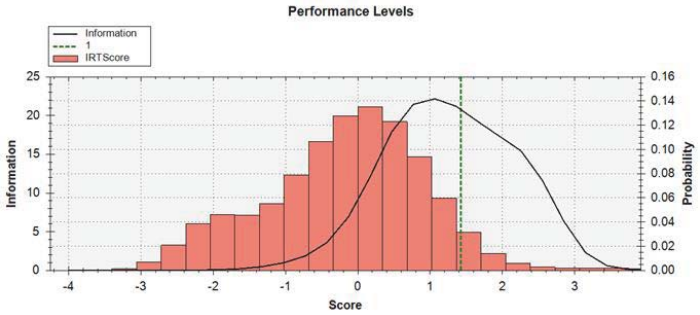

Figure 4: Performance level for the Pre-Test assessment of Community Development Workers.

The response probability setting is set to default of 67 as it is the most common practice to see what is tends to be statistically optimal at the item level. In general, considering the following thresholds as stated by Nuzzo,R.(2014): that if correlation is significant at <0.01 level (2-tailed) then it is very strong assumption against neutral hypothesis, while if correlation is significant at $0.01 < p < 0.05$ then it is strong assumption against neutral hypothesis, and if the correlation is significant at $0.05 < p < 0.1$ it is low assumption against neutral hypothesis while if the Correlation is significant at $p > 0.1$ there is no assumption against the neutral hypothesis. This Pre-Test assessment result has a very strong assumption against the neutral hypothesis as the Correlation is significant at the 0.01 level (**. 2-tailed).

9.2 Post-Assessment Analysis Result for Self-Assessment of Community Development Workers on ICT Skills

Table 4 shows the Post-Test Analysis Result for Self-Assessment of Community Development Workers on ICT Skills. For item 13 in table 4 of the assessment, "*To what extent can you use a Really Simple Syndication (RSS) feeds reader to collect information?*" the highest result was produced in NO SKILLED category with value of 13.8%, indicating that less that 13% of the participants did not increase in knowledge or no skill in the rest of the items and also produce the lowest result in a GOOD SKILLED category with value of 20.6. Table 4 also shows that item 12 "To what extent can you use a government portal/website to access and download government agency related information and forms?" produced the lowest result in LIMITED SKILLED with value of 7.9% and Item 11 "To what extent can you use social media (sms, email, twitter) to facilitate mobile interaction between community members and government representatives?" also produced the lowest result in AVERAGE SKILLED category with value of 20.6%, while item 9 "To what extent can you use a search engine (search engine in a government portal/website) to find funding opportunities, donors and development agencies?"

produced the lowest result in EXPERT SKILLED category with value of 5.8%.

Table 4: Post-Assessment Response Frequency for Self-Assessment of Community Development Workers on ICT Skills.

Item Description		Valid						Missing	Total
		No Skill	Limited Skill	Average Skill	Good Skill	Expert Skill	Total		
ITEM 2 (Word Processor)	Freq.	11	24	49	87	16	187	2	189
	%	5.8	12.7	25.9	46.0	8.5	98.9	1.1	100.0
ITEM 3 Use Template	Freq.	9	18	44	95	21	187	2	189
	%	4.8	9.5	23.3	50.3	11.1	98.9	1.1	100.0
ITEM 4 (Use power Point)	Freq.	12	19	64	77	14	186	3	189
	%	6.3	10.1	33.9	40.7	7.4	98.4	1.6	100.0
ITEM 5 (Presentation Software)	Freq.	13	25	70	63	16	187	2	189
	%	6.9	13.2	37.0	33.3	8.5	98.9	1.1	100.0
ITEM 6 (Search Engine For Crop Management)	Freq.	15	21	68	62	21	187	2	189
	%	7.9	11.1	36.0	32.8	11.1	98.9	1.1	100.0
ITEM 7 (Search Engine For Crop Marked)	Freq.	17	28	62	65	13	185	4	189
	%	9.0	14.8	32.8	34.4	6.9	97.9	2.1	100.0
ITEM 8 Search Engine For Supply Chain)	Freq.	19	28	63	60	17	187	2	189
	%	10.1	14.8	33.3	31.7	9.0	98.9	1.1	100.0
ITEM 9 Search Engine For Dev. Agencies)	Freq.	21	16	68	71	11	187	2	189
	%	11.1	8.5	36.0	37.6	**5.8**	98.9	1.1	100.0
ITEM 10 (Electronic spreadsheet)	Freq.	14	20	55	84	14	187	2	189
	%	7.4	10.6	29.1	44.4	7.4	98.9	1.1	100.0
ITEM 11 (Use Social Medial)	Freq.	7	16	39	88	38	188	1	189
	%	3.7	8.5	**20.6**	46.6	20.1	99.5	.5	100.0
ITEM 12 (Downloading From Website)	Freq.	9	15	44	94	26	188	1	189
	%	4.8	**7.9**	23.3	49.7	13.8	99.5	.5	100.0
ITEM 13 (Use Really Simple Syndication)	Freq.	26	25	79	39	15	184	5	189
	%	**13.8**	13.2	41.8	**20.6**	7.9	97.4	2.6	100.0
ITEM 14 (Use Electronic Media)	Freq.	23	23	68	53	21	188	1	189
	%	12.2	12.2	36.0	28.0	11.1	99.5	.5	100.0
ITEM 15 (Electronic Media To Setup Meeting)	Freq.	13	29	47	78	20	187	2	189
	%	6.9	15.3	24.9	41.3	10.6	98.9	1.1	100.0
ITEM 16 (Setup Group Email Account)	Freq.	16	23	52	68	28	187	2	189
	%	8.5	12.2	27.5	36.0	14.8	98.9	1.1	100.0
ITEM 17 (Use Spreadsheet For Financial Budget)	Freq.	21	29	60	56	21	187	2	189
	%	11.1	15.3	31.7	29.6	11.1	98.9	1.1	100.0
ITEM 18 (Capturing of Data)	Freq.	17	26	67	62	16	188	1	189
	%	9.0	13.8	35.4	32.8	8.5	99.5	.5	100.0
ITEM 19 (Creating new Networking)	Freq.	15	28	61	63	20	187	2	189
	%	7.9	14.8	32.3	33.3	10.6	98.9	1.1	100.0
ITEM 20 (Access Government Information)	Freq.	24	30	66	54	8	182	7	189
	%	12.7	15.9	34.9	28.6	4.2	96.3	3.7	100.0

9.2.1 Post-Test Scale Review Percent score

In percentage score, each bar represents a range of scores, and the height of each bar represents the proportion of examinee with scores at that range. In the chart on the figure 5, column height shows frequency - the number of examinees in overall item score grouping. For example, approximately 33% of CDW received a test score between the ranges score of 62% to70%.

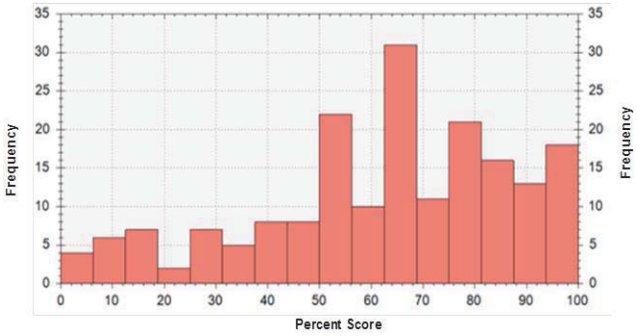

Figure 5: The Percent score for the Post-Test assessment of Community Development Workers on ICT.

9.2.2 Post-Test Performance Standards

In a statistical analysis, the probability value in figure 6 indicates the probability of obtaining same value for a model formulated between two hypotheses, from these hypothesis one is represent as "neutral" (or "null") while the other is placed under the hypothesis measuring. Threshold is set previously and the probability is less than the default threshold (traditionally 5% or 1%). However, we can accept the measure hypothesis as valid and reject the neutral hypothesis. Additionally, the performance standard result for this assessment is valid as the threshold is -0.15, with the mean of -0.53, and the standard deviation is 0.74. However, the probability is less than threshold.

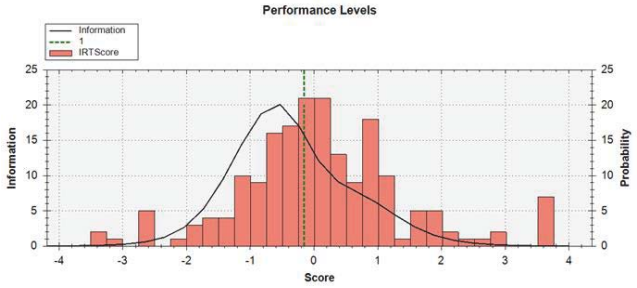

Figure 6: Performance level for the Post-Test assessment of Community Development Workers on ICT.

The response probability setting is set to default of 67 as it is the most common practice to see what is tends to be statistically optimal at the item level. In general, considering the following thresholds as stated by Nuzzo,R.(2014): that if correlation is significant at <0.01 level (2-tailed) then it is very strong assumption against neutral hypothesis, while if correlation is significant at $0.01 < p < 0.05$ then it is strong assumption against neutral hypothesis, and if the correlation is significant at $0.05 < p < 0.1$ it is low assumption against neutral hypothesis while if the Correlation is significant at $p > 0.1$ there is no assumption against the neutral hypothesis. This Post-Test assessment result has a very strong assumption against the neutral hypothesis as the Correlation is significant at the 0.01 level (**. 2-tailed).

10. Conclusion and Discussion

In checking what is the average percentage score for the Pre-Test assessment for all participants, the Pre-assessment percent score for all participant (N=327), was 23.09%, with standard deviation of 21.86 and reliability of 0.95. However the average percentage score for the Post-Test assessment for all participants, (those who completely participated in training program) still needs to be determined. The Post-assessment percent score for the group who completed the training (N=189), the average score is 62.43%, with standard deviation of 24.88 and reliability of 0.94.

Several data driven studies have shown that there is always an improvement between the Pre-Test assessment and the Post-Test assessment. The result from this research shows that there was a significant improvement in the performance of the participants at the Post-Test assessment compared to the Pre-Test assessment. This high performance can be attributed to efficient planning during the training programme. As mentioned with the model, Pre-Test

Assessment involves testing the participants before the commencement of the training programme while Post-Test Assessment involves testing the participants at the end of the training programme. With this, the knowledge, attitudes, or behaviour of the participants are evaluated. As anticipated, most participants in the Post-Test Assessment Stage performed exceptionally well, but it was observed that the participants who failed to perform at the Post-Test Assessment did not complete the Post-Test Training.

These results mean that this experiment not only helped the students understand and gain better skills and knowledge from the training, but also helped the establishment an understanding of and control over the quality of their education output.

11. Limitation of the Study

The limitation of the study is the relatively small sample size of Post-assessment used to investigate the impact of the longitudinal assessment model at the end of the training program. Furthermore, the assessment was only carried out at a single-stage (T_0 to T_1) which is from one point to another. As said earlier, that weakness with the single-stage approach is its inability to distinguish an individual who maintains Pre-teaching knowledge during the course of the session from another individual who forfeits knowledge and then relearns the knowledge again during the knowledge section or term. This limits the authentication as we have just measured the improvement from the assessment of T_0 to T_1. The pedagogy would require ensuring that these items with low knowledge ability in T_0 are addressed and improved by the next assessment of T_1. It is hoped that if the renormalization change continues for more than two-stage (e.g. T_1 to T_2 and T_n) it will improve the knowledge gained more, as the pedagogy will also be redesigned to readdress the situation.

References

Bond, L. 2009. Toward Informative Assessment and a Culture of Evidence: Results from Strengthening Pre-Collegiate Education in Community Colleges. *Carnegie Foundation for the Advancement of Teaching*.

Hake, R. 2007. Possible palliatives for the paralyzing pre/post paranoia that plagues some PEP's. *Journal of MultiDisciplinary Evaluation*, 3 (6): 60-71.

Hake, R. R. 1998. Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics assessment data for introductory physics courses. *American journal of Physics*, 66 (1): 64-74.

Hovland, C. I., Lumsdaine, A. A. and Sheffield, F. D. 1949. Experiments on mass communication.Studies in social psychology in World War II, Vol. 3.

Libarkin, J. C., Anderson, S. W., Deeds, D. and Callen, B. 2006. Development of the geoscience concept inventory. In: *Proceedings of the National STEM Assessment Conference, Washington DC*. 148-158.

McConnell, D. A., Steer, D. N., Owens, K. D., Knott, J. R., Van Horn, S., Borowski, W., Dick, J., Foos, A., Malone, M. and McGrew, H. 2006. Using concepassessments to assess and improve student conceptual understanding in introductory geoscience courses. *Journal of Geoscience Education*, 54 (1): 61.

Meltzer, D. E. 2002. The relationship between mathematics preparation and conceptual knowledge gains in physics: A possible "hidden variable" in diagnostic preassessment scores. *American Journal of Physics*, 70 (12): 1259-1268.

Olugbara, O.O., Millham, R., Heukelman, D., Thakur, S., Wesso, H.W., Sharif, M. 2014. Determining e-Skills Interventions to Improve the Effectiveness of Service Delivery by Community Development Workers. Accepted for Publication by ISI/iNeSI.

Rogosa, D. 1995. Myths and methods:"Myths about longitudinal research" plus supplemental questions. *The analysis of change*, 3: 66.

Ross, S. M. 2009. *Introduction to probability and statistics for engineers and scientists*. Academic Press.

Suskie, L. 2010. *Assessing student knowledge: A common sense guide*. John Wiley & Sons. Place of publication requirred

Willett, J. B. 1989. Some results on reliability for the longitudinal measurement of change: Implications for the design of studies of individual growth. *Educational and psychological measurement*, 49 (3): 587-602.

Willett, J. B. 1997. Measuring change: What individual growth modeling buys you. *Change and development: Issues of theory, method, and application*, 213: 243.

Appendix

(Assessment Items)

Item 1: indicate which of the following tasks you do in your workplace to be performed using ICT
Item 2: To what extent can you use a word processor to create a human settlement report?
Item 3: To what extent can you use a document template to report poverty issues to an information manager?
Item 4: To what extent can you use electronic presentation software such as PowerPoint to create and convey information on HIV/AIDS awareness?

Item 5: To what extent can you use presentation software to create appropriate posters to manage utilities usages, human right and civil responsibility issues?
Item 6: To what extent can you use a search engine (search engine in a government portal/website) to discover crop management information?
Item 7: To what extent can you use a search engine (search engine in a government portal or a website) to access crop market prices?
Item 8: To what extent can you use a search engine (search engine in a government portal or a website) to supply chain tender?
Item 9: To what extent can you use a search engine (search engine in a government portal/website) to find funding opportunities, donors and development agencies?
Item 10: To what extent can you use an electronic spreadsheet to analyse data?
Item 11: To what extent can you use social media (sms, email, twitter) to facilitate mobile interaction between community members and government representatives?
Item 12: To what extent can you use a government portal/website to access and download government agency related information and forms?
Item 13: To what extent can you use a Really Simple Syndication (RSS) feeds reader to collect information?
Item 14: To what extent can you use electronic media to synchronize calendars on mobile and desktop devices?
Item 15: To what extent can you use electronic media to set up a meeting across and within different spheres of government?
Item 16: To what extent can you use electronic media to set up a group on an email account?
Item 17: To what extent can you use an electronic spreadsheet to set up a financial budget?
Item 18: To what extent can you use electronic media to capture associated data in a structured format?
Item 19: To what extent can you use social media to create new networking collaboration?
Item 20: To what extent can you use search engines to access government information and services (health, agriculture, education, funding, aviation, tourism)?