Can Online Delivery Result in Comparable Achievement of Course Outcomes and Student Success in Different Computer Science Courses?

Waleed Farag
Computer Science Department
Indiana University of Pennsylvania
Indiana, PA 15705
farag@iup.edu

Sanwar Ali
Computer Science Department
Indiana University of Pennsylvania
Indiana, PA 15705
sanwar@iup.edu

***Abstract*—This paper investigates reliable answers to the following proposed research questions: Will changing course contents and difficulty levels affect the achievement of course intended learning outcomes and the levels of students' success? Detailed analyses and findings of a proposed course assessment framework that primarily uses direct assessment techniques are presented. Two online courses are compared, an introductory programming class versus a computer literacy one. The paper employs two different data sets. The first set are collected then analyzed to compare the levels of achievement of Intended Learning Outcomes (ILOs) across the studied groups. ILO achievement levels are represented as EAMU vectors then used collectively in the comparison. Statistical analyses of these vectors using both Chi-square and t-tests have shown that there are no statistically significant differences between the two groups. The second data set is used to directly measure and compare a number of factors representing students' success. The obtained results for the majority of success measures denoted that there are no statistically significant differences between the two groups except for withdrawal rate and resource utilization levels which were higher in the introductory programming group. Discussions of the obtained results are provided along with how such findings can support the effectiveness of online delivery.**

Keywords*—measuring course effectiveness, online programming courses, evaluating learning outcomes, measuring student success

I. INTRODUCTION

The ever-increasing use of online course offerings across various disciplines including the Engineering/Computer science fields necessitates the need to better understand this non-traditional delivery mechanism and to ensure its effectiveness and applicability under different circumstances. One of the reasons for the popularity of online delivery is the flexibility that it provides for both the enrolling students and the educational institutions. For students, particularly those that are working full- or part time, it provides a unique opportunity to obtain college credits. On the other hand, the educational institutions see online offerings as an excellent opportunity to reach out to non-traditional students and thus improve recruitment.

In spite of the great opportunities that online delivery provides, it also introduces various challenges [1-2] that need to be addressed for this trend to continue successfully. Some of the issues that are posed as a result of such wide proliferation include:

1. Measuring and improving the effectiveness of online courses.
2. Identifying the best audiences for such delivery mechanism.
3. Researching effective approaches to broaden its applicability and improve retention rates.
4. Studying the ramifications of such wide spread use on higher education and its institutions.

Given such spread, one should keep in mind that online delivery is a unique mechanism for knowledge transfer that has its own characteristics and peculiarities that can be different from those of traditional delivery mechanisms. Therefore, many of the assumptions and findings that generally apply to traditional mechanisms might not apply to online environments. As will be discussed at the end of this section, many researchers have realized the importance of addressing these issues to sustain the success of this trend. In spite of many interesting research studies, this field is still in its infancy and there is a real need for novel approaches and solutions to these outstanding challenges and problems.

The study presented is part of a multi-year, comprehensive research project with the main objective of proposing a reliable framework for evaluating the effectiveness of computer science course delivery. This proposed framework employs both direct assessment and indirect assessment methodologies. In direct assessment techniques, data are derived based on actual student performance in the class. Examples of direct assessment metrics include measuring the achievement levels of course ILOs. On the other hand, indirect assessment techniques use data collected from the student themselves expressing their own perspective of course effectiveness.

This paper addresses the following research question: Will changing course contents and difficulty levels affect the achievement of course intended learning outcomes and the levels of students' success? To provide reliable answers to the posed research question, this paper presents the analyses and findings resulted from applying direct assessment techniques on data sets collected from both groups. Furthermore, the results of using indirect assessment techniques are briefly introduced in Section V.

Our initial research hypothesis can be stated as: Changing the course from an introductory programming course to a computer literacy course will have a noticeable effect on students' achievement of course outcomes and degree of success. This hypothesis was based on observations that some students find introductory programming classes more challenging compared to computer literacy ones which is likely to affect student performance and success. To test the proposed hypothesis, direct assessment techniques are discussed in this paper. The direct assessment metrics were measured by collecting and analyzing two data sets for each group. The first set is used to assess the achievement levels of the ILOs and compare these levels across the two studied groups. The second set is used to measure then compare the degree of students' success and resources utilization in both classes.

A brief literature review of the emergence of online delivery, its recent trends, and examples of research studies focusing on comparing online course effectiveness are presented below. In the past decade, many academicians have delivered various computer science courses online. Two representative examples of these attempts are given in [3-4]. It is also worth noting that one of the noticeable recent trends in online delivery is the introduction of Massively Open Online Course Systems (MOOCs) such as coursera and edX [5-6]. MOOCs trend seems to be promising in spite of challenges identified by some researchers [7]. A positive indicator of the increasing popularity of this trend is the recent offering of the first accredited Master of Science in Computer Science by GIT using MOOC delivery format [8].

Several researchers have studied the evaluation of online courses. Reference [9] presented a recent study to evaluate several delivery modes including online delivery. This study focused on measuring the perception of students regarding the effectiveness of each delivery mode. Results obtained along with traditional student evaluation data were used to improve future offerings. In [10], Helm, Powell and Ice introduced a systematic evaluation process for online courses. This proposed process employed tools developed to assess classroom and homework equivalent contact hours in an online environment. In [11], an interesting quantitative evaluation model was proposed that considered both the interdependence among measures and the inherent noise in subjective perceptions in order to evaluate the effectiveness of uncertain e-learning systems. Another study that evaluated how the number of students in a class affected the quality of the online delivery was presented in [12]. This study found that large class size negatively impacted students' satisfaction as well as limited the quality and quantity of student–student and student–instructor interactions. Reference [13] presented a study that first validated a proposed tool to measure students' evaluation of online courses then used that tool to examine students' evaluations of their online courses. That study found that students rated content organization and format higher than other performance measures. A number of assessment methodologies and best practices for evaluating online offering are also described in [14].

From this brief survey, one can notice that online delivery is certainly on the rise. In addition, several researchers have realized the importance of reliably assessing the effectiveness of such offerings. This study presents a novel proposed assessment framework to evaluate the impact of course contents on the achievement levels of ILOs and students' success.

II. STUDY DETAILS

This work discusses results from a comprehensive experimental research study with the objective of investigating the effect of varying course contents and difficulty levels on the achievement of the intended learning outcomes and other success indicators. The study compares two groups using a proposed framework for assessing the effectiveness of such courses. Both groups are computer science courses that are offered fully online in our department.

The first group data were taken from our Computer Literacy (CS 101) class. This course is an introduction to computers and covers basic topics such as computer hardware, applications and system programs, operating systems, networking, and security. An important laboratory component of this course covers various MS-office applications. This course is usually taken by Computer Science minors as well as business and other majors. The collected data were taken from recently taught sections that were delivered fully online in the last two years.

The second group comprises data obtained from our introductory programming course (CS 110). This course is a traditional introduction to algorithm development and programming. It covers basic programming concepts such as data types, allocation, expressions, I/O and advances to discuss decision making, looping, structured data types, functions, recursion, OOP, and simple GUI concepts. This course is required of all Computer Science majors and is also taken by science and other interested students from various majors. The collected data were taken from several sections of this course that were taught fully online in the last five years.

Although the delivered content in these two courses have different focuses, these fully online courses have similar structures. Both feature comprehensive syllabi that highlight the self-paced nature of the course while clearly expound all policies and expectations. Asynchronous activities in both courses include projects, laboratory environment, online exercises, and several discussion forums. Synchronous activities including chat sessions and phone calls were also lightly used. The similar structures of both studied groups will add to the homogeneity of the study and help exclude extraneous factors. Moreover, in a trial to ensure study reliability, all collected data were obtained from sections that were taught by the same instructor using the same teaching style and Learning Management System (LMS).

As stated before, this research tries to produce a reliable answer to the following question: Will the change in course contents from "Intro to Programming" to "Computer Literacy" have a significant impact on course effectiveness? To answer this question, we used an inclusive framework for measuring the effectiveness of the delivered courses then conducted an in-depth statistical analyses to compare the two studied groups. The employed framework uses both direct and indirect performance assessment factors. Two direct assessment methods are used. The first one measures the achievment levels of the course intended learning outcomes (ILOs) while the second one employs a set of student success metrics. The focus of this paper will be discussing the two sets of direct performance indicators, followed by reporting the details of the statistical analyses and findings. The two indirect assessment methods assess the students' perception of various online course delivery metrics and the degree of students' satisfaction with the course. Results of the indirect assessment experiments are briefly discussed towards the end of this paper while the details of these experiments are the focus of another work. Lastly, as discussed previously, our research hypothesis posed by this study states that the change in course contents will result in differences when measuring and comparing various course assessment factors.

III. ANALYSIS OF COURSE LEARNING OUTCOMES ACHIEVMENT

For CS 101, various course objectives can be summarized into four main course outcomes (listed below) that each student completing this course must attain at a certain level.

1. Demonstrate deep understanding of essential computing concepts and how to have a successful digital life style (O1).
2. Illustrate the ability to use various MS-office applications simulated labs (O2).
3. Apply knowledge of various MS-office applications to generate and edit professional office documents (O3).
4. Recognize the importance of actively participating in class discussion forums and other online activities (O4).

Similarly, CS 110 has four main Intended Learning Outcomes (ILOs) given below.

1. Put into practice effective use of an Integrated Development Environment to edit, compile, and run programs (O1).
2. Demonstrate the ability to develop algorithms from problem specifications and apply various structured programming techniques to proficiently transform them into programming code (O2).
3. Illustrate the ability to use debugging and testing techniques to locate and fix errors to ensure program correctness (O3).
4. Recognize the proper use of the language's constructs and apply this knowledge in creating effective programs (O4).

To assess the level of attainment of each ILO, a set of course graded activities are selected as a measure for this particular outcome. Each student's achievement in this set is measured to compute a percent value indicating how well a student attains each outcome. Then, a number of thresholds were proposed to differentiate Exemplary (E), Adequate (A), Minimal (M), and Unsatisfactory (U) performance. This procedure is applied to the four outcomes and all students in each course to develop an EAMU vector for each ILOs. The EAMU notation was introduced in [15] and has been adopted in previous research studies [16-17]. As an example, the obtained overall EAMU vector for outcome 1 (O1) in CS 110 offerings was [49 9 4 4]. This means that out of the 66 students who passed the course, 49 students exhibited exemplary attainment of O1, nine attained adequate levels, four measured at the minimal level and the performance of the remining four students was unsatisfactory. The computed EAMU vectors for each outcome along with the percentage of each level are listed in Table 1 for both groups.

The results in Table 1 show somewhat comparable achievement levels across the two groups. For instance, in outcome 4, the percentages of students who achieved Exemplary (E) performance are 35% and 38% in the 110 and 101 classes, respectively. Table 1 also indicates that the sample size in the 110 group is 66 students, and 56 students for the 101 group. Where: N is the sample size, EAMU vectors are the computed EAMU vectors for each outcome and the last four columns are for the percentages of the exemplary, adequate, marginal and unsatisfactory performances respectively.

TABLE I. OVERALL EAMU VECTORS AND PERCENTAGES

Outcome	N	EAMU vectors	E%	A%	M%	U%
O1-110	66	[49 9 4 4]	74%	14%	6%	6%
O2-110	66	[35 17 9 5]	53%	26%	14%	8%
O3-110	66	[40 13 4 9]	61%	20%	6%	14%
O4-110	66	[23 29 10 4]	35%	44%	15%	6%
O1-101	56	[17 32 6 1]	30%	57%	11%	2%
O2-101	56	[44 5 4 3]	79%	9%	7%	5%
O3-101	56	[14 20 14 8]	25%	36%	25%	14%
O4-101	56	[21 26 6 3]	38%	46%	11%	5%

Given that the objectives of the two courses are not identical, we computed the average of each achievement level over all four studied ILOs and used the overall averages as a general measure of the degree of achievement of the studied course outcomes. These averages were compared across both courses. For example in the 110 class, the average of achieving Exemplary level over all four ILOs is ((49+35+40+23)/4) which equals 37. Applying the same procedure on the data in Table 1 gives us the overall average EAMU vector for 110 as [37 17 7 5] and the corresponding vector for 101 as [24 21 7 4]. The percentages for achieving each components of the EAMU vectors will then be [56% 26% 11% 8%] and [43%

38% 13% 7%] for the 110 and 101 classes, respectively. When comparing the collective achievement of the overall class ILOs (using the overall EAMU percentages just computed), we notice that the percentage of students achieving Exemplary performance in 110 is higher than the corresponding percentage for students in the 101 class by 13%. For the Adequate performance level, the reverse is true in which the 101 group has a higher percentage than the corresponding one for the 110 group by 12%. Nevertheless, comparable performance can be noticed if we combine both the E and A levels in both cases (82% for 110 and 81% for 101). Similarly, both M and U levels have shown comparable percentages in both groups.

From the percentages in Table 1, one can initially conclude that there were generally comparable achievements of ILOs in both classes and the contents and degree of difficulty of the delivered materials did not have a considerable impact on the achievement of ILOs for those students who completed the class. In order to ensure the statistical correctness of this initial observation, additional analyses need to be done. The first proposed statistical procedure treats the data as Categorical and thus uses the Chi-Square test to compare both groups. The obtained Chi-square test results can be stated as:

> For achieving unsatisfactory (U) performance level: There was NO significant association between the course contents (110 or 101), $\chi 2$ (3) = 2.5, p = .48. It is also worth noting that the computed odds ratio, 1.07, indicates that the odds of students achieving U level in the online 110 class were 1.07 times if they were taking the online 101 class which supports the insignificant Chi-square test result.

Another statistical test was also employed to confirm the results obtained by the Chi-square test, above. In this experiment, the data are treated as continuous ones. To do that, each EAMU vector for each outcome (of a certain offering) is transformed into a scalar value. The transformation is accomplished by using a proposed weighted sum formula (see equation 1) that places increasingly higher weights on good performance (A and E) while penalizes marginal and unsatisfactory performance by assigning them diminished weight values. The formula is normalized to the number of students (N) in the offering to produce a comparable scale. For each course offering, a scalar value was computed that represents achievement of a specific ILO. These are then averaged over the studied four ILOs (a process similar to the one described in the Chi-square analyses above) to determine one overall scalar value for each offering that expresses the overall degree of achieving the ILOs in this specific offering. These data are then analyzed using Shapiro-Wilk and Levene's to check whether the data conform to the parametric test assumptions or not. That analysis revealed that the data are normally distributed within each group. In addition, the homogeneity of variance (Levene's) test produced insignificant results where the statistical test value for the Levene's test (F) was .23 and the test significance (p values) was .65. Given that the data conform to the parametric test assumptions, the Independent Samples t-Test was used to compare the achievement levels for the overall course outcomes in both cases. The results of using the t-test are presented in Table 2, where μ is the mean value, SE is the standard error, t is the test statistical value, df is the degree of freedom, and r is the effect size which provides a measure of the importance of the effect [18].

$$\text{Outcome(i)} = \frac{1}{N}(0.55E + 0.25A + 0.15M + 0.05U) \tag{1}$$

TABLE II. RESULTS OF T-TEST FOR OVERALL ILOS COMPARISON

Outcome overall	**Descriptive values**		**t-test**			
	μ	***SE***	***t***	***df***	***p***	***r***
Overall ILOs-110	.40	.01	1.59	8	.15	.49
Overall ILOs-101	.35	.04				

In Table 2, the t-test significance (p value) is greater than .05 which indicates that both groups exhibit similar distributions. Thus, the results of the t-test for the overall achievements of ILOs can be reported as follow:

> On average, students taking the online 110 class scored slightly higher in achieving overall ILOs (μ = 0.396, SE = 0.013) than students taking the online 101 class (μ = 0.346, SE = 0.036). This difference was NOT significant t(8) = 1.587, p > 0.05. Moreover, the effect size (r) is 0.49.

In summary, the results of both tests (the Chi-square test and the t-test) confirmed that both groups have similar distributions for the achievement levels of the studied ILOs. In other words, there were no statistically significant differences between the distributions of the studied groups. Thus, one can conclude that the change in course contents from 110 to 101 and the varying levels of materials difficulty have no significant effect on the achievement of the Intended Learning Outcomes in both courses. These results do not support our main research hypothesis in this study.

IV. EVALUATING STUDENTS' SUCCESS AND RESOURCES UTILIZATION

The second proposed set of direct assessment performance indicators focused on the degree of students' success and resources utilization. The chances of students receiving the highest grade (A), failing grade (F), or withdrawing from the course (W) are compared as detailed in the following sub-section.

A. *Chances of Receiving A, F, and W*

The three factors analyzed in this sub-section can be defined as follow:

- Chances of obtaining the highest grade (A-Factor): number of students who received A grade to the total number of students who passed the class.

- Chances of obtaining a failing grade (F-Factor): number of students who received F grade to the total number of students who completed the class.
- Chances of withdrawing from the course (W-Factor): number of students who withdrew from the class to the total number of students who enrolled in the class.

The data for these three factors are analyzed as categorical data. To analyze these data, Chi-Square ($\chi 2$) test was used and the results are given in Table 3 where # is the number of students achieving this level and N is the overall sample size. For example, Table 3 shows that the percentages of students who received A grade are 47% and 34% in the 110 and 101, respectively. $\chi 2$ tests for the A and F factors produced insignificant result ($p > .05$). For example, the obtained results for A-Factor can formally be reported as:

> There was NO significant association between the class contents (110 or 101) and whether or not students can receive a final grade of A in the class $\chi 2$ (3) = 3.57, p = .31. It is also worth noting that the computed odds ratio, 1.75, indicates that the odds of students receiving A in the online 110 class were 1.75 times if they were taking the 101.

TABLE III. CHI-SQUARE TEST RESULTS FOR THREE SUCCESS FACTORS

Success Factor	#	N	%	Chi-Square test results			
				χ2	*df*	*p*	*odds ratio*
A-110	31	66	47%	3.57	3	.31	1.75
A-101	19	56	34%				
F-110	36	102	35%	7.70	4	.10	2.04
F-101	15	71	21%				
W-110	69	171	40%	21.55	1	.00	5.23
W-101	9	80	11%				

On the other hand, the obtained results in Table 3 for the W-factor gave significant test statistics. The results for the withdrawal factor can be stated as follow:

> There was a significant association between the class contents (110 or 101) and whether or not students withdrew from the class $\chi 2$ (1) = 21.55, $p < .05$. This result is also supported by the computed odds ratio, 5.23, indicating that the odds of withdrawal in the online 110 class were 5.23 times if students were taking the 101 class.

The results obtained above are not entirely surprising for two reasons. First, high withdrawal rates have frequently been reported in various online courses [19]. Second, it is believed that the relatively difficult course contents presented in 110 have led to this statistically significant association between the course type and the withdrawal rate.

B. *Analysis of Overall Percentage Grade*

This analysis focuses on comparing the average overall percentage grade in the 110 course with the corresponding percentage for the 101 group. Such a comparison gives us better understanding of whether the course content and level of difficulties have affected the average overall percentage grade for all students in the class. The research hypothesis in this analysis states that the distribution of final grade percentages will be different across categories of class type. Before selecting a test statistic, the data were first checked for conformance with parametric assumption as done before. The Shapiro-Wilk test results for 101 class group indicates normally distributed data but the test gives significant result ($p = 0.02$) for the 110 data. Fig. 1 depicts the non-normal frequency distribution of the 110 class grades with apparent negative skewness and kurtosis. Please note that there is a chance that some final percentages may exceed 100% with the availability of extra credits. In addition, the Levene's test also gives significant results indicating that the variance of the two groups is not homogenous. Since the data do not satisfy the parametric test assumptions, we adopted the Mann-Whitney, a non-parametric test, to compare the averages in both groups. This test yielded a non-significant difference, a result that does not support the initial research hypothesis given above. To summarize the test outcome:

> Students final grades in the online 110 class (Median = 89.28) did not differ significantly from those enrolled in the online 101 class (Median = 85.75), U = 1574.00, z = -1.41, ns: $p > 0.05$, r = -0.13 which represents a small effect. This was also apparent in the close values of the average final grades in the 110 and 101 classes, 85.67 and 84.04, respectively.

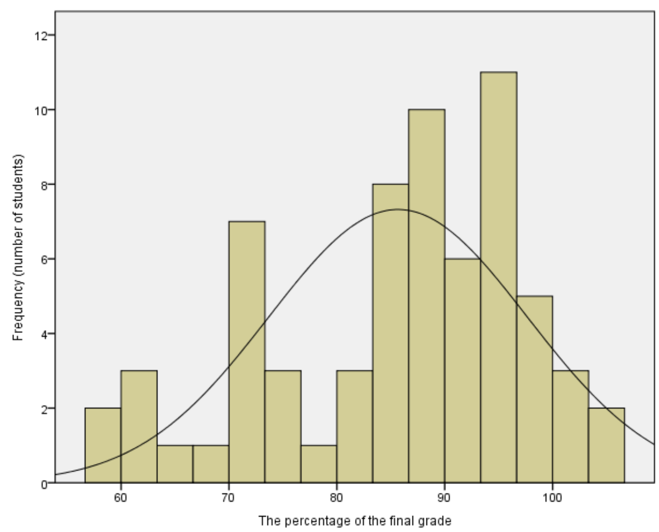

Fig. 1. Non-normal distribution of grades in 110 course

C. *Comparing Resources Utilization*

In this subsection, we use the frequency of utilization of course online resources as a rough indicator to the involvement of the students, then compare the computed values across both of the studies groups. Our research hypothesis states that students enrolled in the 110 class will utilize the online class contents more than those enrolled in the 101 due to differences in course contents and difficulty of the covered materials. As expected students in the 110 online class have higher access frequency to course online materials (μ = 1248) while the corresponding mean for the 101 students is (μ = 538). The Shapiro-Wilk test reveals that the data is normally distributed for the 101 group but not in the 110, see Fig. 2 that depicts the histogram distribution of frequency of access to various online course contents for the 101 group. Data in Fig. 2 shows a typical normal distribution with a positive kurtosis (0.56) and slightly positive skewness (0.09). Besides, the Levene's test found that the variance is not homogeneous; therefore, Mann-Whitney, a non-parametric test, was used yielding the following:

> Students frequency of access to course contents in the online 110 class (Median = 1235) differs significantly from those enrolled in the online 101 class (Median = 510), $U = 1523.50$, $z = -6.23$, $p < 0.05$, $r = -0.48$ which represents a large effect that accounts for about 25% of the variance.

It is worth noting that the used online contents in both classes share large similarities in content structure and access requirements. This finding supports our research hypothesis mentioned above. Thus, one might conclude that the difficult nature of the programming course content (110) encouraged students to utilize more frequently the online resources compared to the other group.

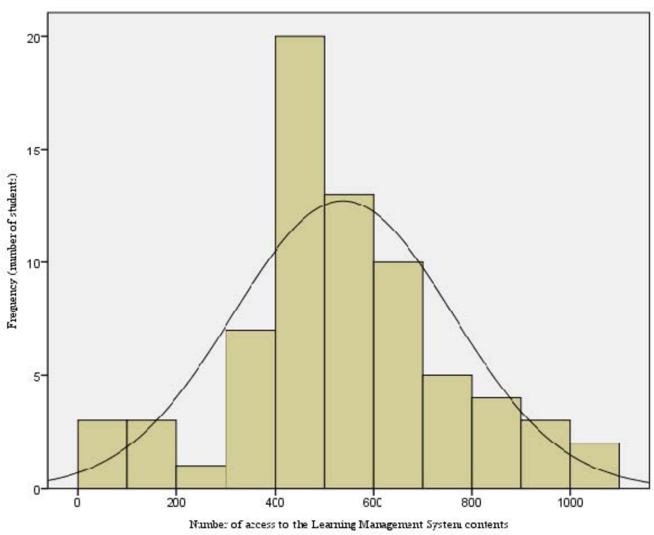

Fig. 2. Normally distributed resource access data in the 101 course.

V. INDIRECT ASSESSMENT RESULTS

As discussed earlier, the goal of this research focused on discussing the direct assessment (mainly the achievement of the intended learning outcomes and the evaluation of several essential success indicators) analyses and results to shed some light on answers to the proposed main research question. The results discussed here represent a subset of the results and findings of a comprehensive quantitative study to evaluate online course effectiveness, using a multitude of performance factors applied to various data sets and scenarios. Because the results of indirect assessment performance metrics have not been discussed in this paper due to space limitation, this section briefly comments on them. Two data sets were used in the indirect assessment. The first one was used to assess the perception of students via a number of online specific performance metrics listed below, most of them are part of the COLLES model [20]:

1. Relevance: How relevant is online learning to the students and their professional practices?
2. Reflection: Does online learning stimulate students' critical reflective thinking?
3. Interactivity: To what extent do students engage online in rich educative dialogue?
4. Tutor Support: How well do tutors enable students to participate in online learning?
5. Peer Support: Is sensitive and encouraging support provided online by fellow students?
6. Interpretation: Do students and tutors make good sense of each other's online communications?
7. Course elements: A combined performance measure that assess various course components such as labs, quizzes, etc.

The majority of studied factors in this data set (5 out of 7) showed non-significant differences between the two groups but the remaining two factors (interactivity and peer-support) exhibited statistically significant results. The second indirect data set was used to compare students' satisfaction levels when taking the 110 class with the same levels when taking the 101 class. The obtained statistical results indicated that there were no significant differences between the two groups in both of the adopted performance measures (course satisfaction and instructor satisfaction). The detailed analyses for all indirect assessment factors are the focus of another work.

VI. CONCLUSIONS

This work presented the analyses and results of a study with the objective of answering the research question: Will changing course contents and difficulty levels affect the achievement of course intended learning outcomes and the levels of students' success? The paper devised the use of a framework in order to achieve its objectives. The framework employed sets of direct and indirect assessment measures to be used to assess course effectiveness. The paper focused on reporting the results of the direct assessment measures. Results from the indirect assessment experiments were briefly discussed in the previous section.

Two groups were the focus of this experimental research study. The data for the first group were collected from several recent online offerings of an introductory programming course while the data for the second group were collected from recent online offerings of a computer literacy course. The paper employed two data sets for each of the studied groups. The first data set was used to assess and compare, across the studied groups, the achievement levels of the Intended Learning Outcomes (ILOs) in each course. Equivalent ILOs for both groups signify that both student populations have acquired comparable levels of knowledge and skills after the completion of the course. In our experiments, the four major ILOs in each course were measured by computing the EAMU vectors for each outcome. The statistical analyses denoted that there were no statistically significant differences between the achievement levels of course Intended Learning Outcomes across the two studied groups. Moreover, a second direct assessment data set of performance measures focused on student success criteria was employed. Data analyses for these factors yielded comparable distributions in most areas. Two exceptions were found. The first was the rate at which students have withdrawn from class, in which significant association between the class type/difficulty levels and the withdrawal rate was present. The second was the resource utilization factor in which there was a statistically significant association between the frequency at which students access the online course resources and the course difficulty.

In conclusion, our main research hypothesis in this study has not been supported by the overall obtained results. It is believed that the interesting findings of this study can strongly support the argument regarding the effectiveness of online delivery. We have obtained comparable achievement levels for course learning outcomes in two courses that are qualitatively different and of varying course difficulty. Similar comparable levels of performance were also obtained for most of the parameters used to measure student success and resources utilization. Such findings attest to the ability of online courses to effectively deliver numerous course contents irrespective of the technical depth of the covered materials.

Lastly, we would like to highlight some strengths and weaknesses of this study. Our sample sizes were reasonably large and have reached 170+ in some experiments. These ranges produced results that are reliable. On the contrary, we have relatively small sample sizes (N = 25) in certain experiments, mainly when measuring one of the indirect assessment factors (students satisfaction). Such sample size was dictated by the fact that participation in these surveys is totally voluntarily (a requirement of our institution IRB approval) in all the sections used in this study. Lastly, the obtained results and the drawn conclusions were mainly for the studied cases (110 and 101). Caution should be exercised when trying to generalize our conclusions on other situations without further investigation.

REFERENCES

[1] A. Azemi, M. Bodek, and G. Chinn, "Teaching an introductory programming course using hybrid e-learning approach," proceeedings of IEEE Frontier in Higher Education (FIE) Conference, 2013, pp. 1911-1913.DOI= http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6685168.

[2] W. Armitage, N. Boyer, S. Langevin, and A. Gaspar, "Rapid conversion of an IT degree program to online delivery: impact, problems, solutions and challenges," proceedings of the SIGITE Conference on information technology education, 2009. pp. 100-107. DOI= http://doi.acm.org/10.1145/1631728.1631758.

[3] E. El-Sheikh, "Techniques for engaging students in an online computer programming course," J. of Systemics, Cybernetics and Informatics, vol. 7, no. 1, pp. 1-12, 2009.

[4] A. Azemi, "An online introductory computer programming course using Matlab: Design steps and what to expect," proceeedings of IEEE Frontier in Higher Education (FIE) Conference, 2015, pp. 1-4, DOI= http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7344186.

[5] Coursera Site 2016. Retrieved April 24, 2016 from https://www.coursera.org/.

[6] EdX Site 2016. Retrieved April 24, 2016 from https://www.edx.org/

[7] L. Chamberlin, and T. Parish, "MOOCs: Massive Open Online Courses or Massive and Often Obtuse Courses?" ACM eLearn, vol. 8, Article No. 1, 2011. DOI= http://dx.doi.org/10.1145/2016016.2016017.

[8] GIT MS in Computer Science Site 2016. Retrieved April 24, 2016 from http://www.omscs.gatech.edu/

[9] J. Mullen and J. Sullivan, "Student-perceived effectiveness of online content delivery modes," proceeedings of IEEE Frontier in Higher Education (FIE) Conference, 2015, pp. 1-4.

[10] J. Helm, K. Powell, and P. Ice, "Evaluating online course quality for student learning and success," proceedings of the World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education, Hawaii, 2011.

[11] M. Tseng, R. Lin, and H. Chen, "Evaluating the effectiveness of e-learning system in uncertainty," J. of Industrial Management & Data Systems, vol. 111, no. 6, pp. 869-889, 2011. DOI= http://dx.doi.org/10.1108/02635571111144955

[12] V. Russell and W. Curtis, "Comparing a large- and small-scale online language course: An examination of teacher and learner perceptions," J. of the Internet and Higher Education, vol. 16, pp. 1–13, 2013. DOI= 10.1016/j.iheduc.2012.07.002

[13] T. Rothman, L. Romeo, M. Brennan, and D. Mitchell, "Criteria for assessing student satisfaction with online courses," J. International Journal for e-Learning Security, vol. 1, 1/2, March/June, pp. 27-32, 2011.

[14] T. Tobin, B. Mandernach, and A. Taylor, Evaluating Online Teaching: Implementing Best Practices, Wiley, 2015.

[15] J. Estell, "Streamlining the assessment process with the faculty course assessment report," Workshop in the 37th IEEE ASEE/IEEE Frontiers in Education Conference, Milwaukee, 2007, pp. W1A-1. DOI= http://doi.ieeecomputersociety.org/10.1109/FIE.2007.4418236

[16] W. Farag, "Comparing achievement of intended learning outcomes in online programming classes with blended offerings," proceedings of the SIGITE Conference on information technology education, Calgary, Canada, 2012 pp. 25-30. DOI= http://doi.acm.org/10.1145/2380552.2380561.

[17] W. Farag, S. Ali, and D. Deb, "Does language choice influence the effectiveness of online introductory programming courses?" proceedings of the SIGITE Conference on information technology education, Orlando, Florida, 2013, pp. 165-170. DOI= http://dx.doi.org/10.1145/2512276.2512293.

[18] A. Field, Discover Statistics Using SPSS. SAGE Publications Ltd, London, UK, .2009.

[19] B. Hare, Classroom interventions to reduce failure & withdrawal in CS1: a field report, J. of Computing Sciences in Colleges, vol. 28, no. 5, pp. 228-235, 2013.

[20] P. Taylor, and D. Maor, "Assessing the efficacy of online teaching with the Constructivist On-Line Learning Environment Survey," proceedings of the 9th Annual Teaching Learning Forum, Perth, Australia, 2000.