An Integrated Model of The Relation Between E-Service Quality and User Satisfaction in IHL

Iham Tariq Ismail Al-Nuaimi¹, Tameem Mohammed Mahmood², Othman Waleed Khalid³, Ammar K Alazzawi⁴

¹Department of Computer Engineering Techniques, Al-Yarmouk University College, Dayala, Iraq
²Department of Computer Engineering Techniques, Al-Yarmouk University College, Dayala, Iraq
³Department of Computer Engineering Techniques, Al-Yarmouk University College, Dayala, Iraq
⁴Department of Computer and Information Science, University Technology Petronas, Seri Iskandar, Malaysia

Ayhemtariq@al-yarmok.edu.iq ,

Abstract. Quality of e-service is one of the critical factors that decide the success or failure of organizations. It may increase competitive advantages as well as enhance the relationships with the customers. Achieving high e-service quality and user satisfaction are challenging since they depend fundamentally on user perception and expectation which can be tricky at times. To date, there is no agreement as to what service quality is, and how it should be measured, whether it is a function of statistical measures of quality including physical defects or managerial judgment, or it is a function of customer perception about the services. This paper deeply dived the quality of e-services offered by five Malaysian Institutes of High Learning (IHL) including two private and three public universities. A quantitative approach was utilized to collect the data and AMOS 21 was used to analyze the data and develop the measurement model. The paper aims to find the relationship between e-service quality dimensions and the user satisfaction by using Conformity Factor Analysis (CFA) test with 320 students in the target universities. The research results indicated that the measurement model has acceptable values and the Structural Equation Model (SEM) has positive impact on the relationship between e-SQ dimensions and user satisfaction. Achieving high user satisfaction can enhance the competitive advantage of the universities in their respective target markets.

Keywords: E-service quality; CFA; Users’ satisfaction; Expectancy Disconfirmation Theory; AMOS

1. INTRODUCTION
Malaysian Institutes of High Learning (IHL) are constantly seeking to use modern technologies in various fields to enhance their services for the dissemination of knowledge. In demonstrating their commitments in academic excellence, research excellence, and professional services, they are subjected to continuous pressure to surpass their competitors in presenting better e-services quality to improve users relationship and to meet their satisfaction [1, 2]. The success of providing their e-services depends significantly on the quality perceived by users on universities portals that offer various services for students. Although some of these services are similar in content, they vary in dimensions such as the look and feel, viewing method, and quality of service. The quality of services in the educational institutions’ portal has received a great attention by researchers [3], because studies have shown that it can increase current users’ retention, attract new users, and hence enhance the online competitive advantages of companies [4]. Attaining high e-service quality and user satisfaction is difficult because they depend on user perception which can’t be measured easily. To date, it is still
inconclusive as to what service quality is and how it should be measured [5]. As such there is a need for a scale to measure e-SQ in higher education sector in order to enhance its e-SQ.

This paper aims to determine the relationship between e-service quality dimensions and the user satisfaction based on expectancy disconfirmation theory model. Conformity Factor Analysis (CFA) test was conducted to validate and evaluate the data collected by a questionnaire with 320 students in five Malaysian Institutes of High Learning (IHL) including two private and three public universities. The CFA is a statistical technique allows explaining relationships among observed variables and with their underlying latent constructs exists. AMOS 21 was used for the analysis, which started from conducting the measurement model to find the relationship between e-SQ dimensions and user satisfaction, and then to come out with the final model.

The paper is organized as follows. Section two explains the expectancy disconfirmation theory model (EDT), while section three presents the significant dimensions of e-SQ in Malaysian IHL. Section four shows the data analysis by conducting SPSS 20 and AMOS 21. Finally, section five illustrates the conclusion and future work of the research.

2. EXPECTANCY DISCONFIRMATION THEORY

Expectancy Disconfirmation Theory (EDT) is based on Cognitive Dissonance Theory (CDT) that was introduced in 1957 by Leon Festinger. The EDT can measure the customer’s satisfaction from the difference between customer’s expectation and experience in perceived products or services [6]. According to EDT, consumers compare their expectations to perceived performance. If perceived performance exceeds expectations, positive disconfirmation results and satisfaction occurs. If expectations exceed perceived performance then negative disconfirmation results and dissatisfaction occurs [7].

Researcher defined different models based on EDT. Oliver (1980 cited in Picazo-Vela, 2009 [7]) proposed a model with positive relationship between expectations and satisfaction, with high expectations leading to high satisfaction as shown in Figure 1.

![Figure 1](image-url)

**Figure 1.** Proposed model with positive relationship between expectations and satisfaction [7]

Bakri and Elkhani proposed the model illustrated in Figure 2 the model includes the same components of the complete model but with different relations [6].

![Figure 2](image-url)

**Figure 2.** Proposed model with positive relationship between perceived performance and satisfaction [6]

Picazo-Vela presented his model as shown in Figure 3 to analyze the impact of online reviews on customer satisfaction.
It is obviously that researchers utilized EDT in different models based on their field of study to investigate the relations between their EDT models components. Based on that and on Premkuma and Bhattacherje, this research utilizes a comprehensive model of EDT that studies the relation between both perceived performance and expectations with satisfaction as shown in Figure 4 [8].

3. SIGNIFICANT DIMENSIONS OF E-SQ IN MALAYSIAN IHL
The significant dimensions of e-SQ in Malaysian IHL have been adopted based on the results of EFA (Exploratory Factor Analysis) [9]. Based on the expectancy disconfirmation theory, these dimensions construct the hypothesized conceptual model for the relation between e-service quality and user satisfaction as shown in Figure 5.
The all constructs of the model are demonstrated below with their statements and hypotheses.

1. **Fulfillment**: Refers to the degree of the website in keeping the promises to provide the required and expected services at reasonable time.
   - All services indicated on the site are available.
   - The website delivers services as expected.
   - The time for service processing is acceptable.
   - The website provides services at real time.
   - It is easy and fast to load pages.
   - The website delivers the required services within the expected time frame.
   - Correct information is provided on each service.
   - The site requires a minimal amount of information to be input by the users.
   - The site offers the necessity services required by the user.

   H4A: There is a positive relationship between fulfillment and expected outcome.
   H4B: There is a positive relationship between fulfillment and perceived performance.

2. **Interactivity**: This refers to the website interaction with the users, and how it encourages them to use the available services, and keeps them on touch.
   The website motivates users to use online services.
   - The website provides channels for communication with the users.
   - The website is visually appealing.
   - The website reminds users prior to deadlines of services.
• The website provides its contents in different languages.
• The website respects the other cultures.
• The website informs users when new services are available.
• The website is well structured.
• It is easy to log into the website.
• The website design is simple and clear.

H7A: There is a positive relationship between interactivity and expected outcome.
H7B: There is a positive relationship between interactivity and perceived performance.

3. Information: Refers to the quality and relevance of provided information to the user through the website.

• The provided information is easy to understand.
• It is easy to navigate through the website to find the required information.
• Information about all services is available.
• The information is always detailed and relevant.

H8A: There is a positive relationship between information and expected outcome.
H8B: There is a positive relationship between information and perceived performance.

4. System Availability: Refers to the proper technical performance of the site, which includes its continuous availability, URL validation, website integrity, easy to lunch and run the website.

• The website is available 24/7.
• The site launches and runs at acceptable response time.
• Site failures are rare.
• URL links work correctly.

H2A: There is a positive relationship between system availability and expected outcome.
H2B: There is a positive relationship between system availability and perceived performance.

5. Reliability: Refers to the performance integrity and the service provider ability to supply accurate service and alternative option in case of service halt.

• The same problems occur frequently.
• Problems in e-services are solved quickly.
• Alternative options are given for each service when the service is disrupted.

H5A: There is a positive relationship between reliability and expected outcome.
H5B: There is a positive relationship between reliability and perceived performance.

6. Privacy/Security: Refers to the secured user personal information and website safety against threats by developing policies to regulate and ensure these issues.

• The website offers high security for personal information.
• The website posts a clear privacy policy.
• The site is safe and free from attacks or potential threats.

H3A: There is a positive relationship between privacy/security and expected outcome.
H3B: There is a positive relationship between privacy/security and perceived performance.

Expected Outcome: Refers to the level of the expectancy of the services provided.

H10: There is a negative relationship between expected outcome and disconfirmation.
H13: There is a positive relationship between expected outcome and user satisfaction.

Perceived Performance: Refers to the level of perceived services.

H11: There is a positive relationship between perceived performance and disconfirmation.
H14: There is a positive relationship between perceived performance and user satisfaction.

Disconfirmation: Refers to comparison between initial expectations and performance.

H12: There is a positive relationship between disconfirmation and user satisfaction.

User Satisfaction: Refers to the level of making the user satisfy with the services provided.

4. DATA ANALYSIS

4.1 Descriptive Analysis
The research aims to measure students’ satisfaction by using AMOS 21 to get the measurement model/Confirmatory Factor Analysis (CFA). Five Malaysian Institute of High Learning (IHL) have been chosen to conduct the survey namely UTP, UM, UKM, UTM, and UNITEN. Around 350 questionnaires (70 for each university) were distributed among the students whom selected based on a simple random sampling. The valid questionnaires were 320 distributed between the five universities as shown in Table 1.

| University Name | Frequency | Percent | Cumulative Percent |
|-----------------|-----------|---------|--------------------|
| Valid           | 69        | 21.6    | 21.6               |
| UTP             | 57        | 17.8    | 39.4               |
| UM              | 63        | 19.7    | 59.1               |
| UKM             | 63        | 19.7    | 78.8               |
| UTM             | 68        | 21.3    | 100.0              |
| UNITEN          | 68        | 21.3    | 100.0              |
| Total           | 320       | 100.0   |                     |

Tables 2 and 3 show the respondents gender and ages. The respondents’ males’ percentage is 58.2% and the female percentage is 41.3%, with an age range from 18 to 51. The largest age group of respondents was between 20 and 30 years old (68.8%) and the smallest age group was more than 51 years old (0.6%).

| Gender | Frequency | Percent | Cumulative Percent |
|--------|-----------|---------|--------------------|
| MALE   | 188       | 58.8    | 58.8               |
| FEMALE | 132       | 41.3    | 100.0              |
| Total  | 320       | 100.0   |                     |

Table 3. Respondents ages
Table 4 illustrates the education levels of the respondents. It indicates that the largest educational group was bachelor degree (50.9%) followed by master (32.8%) and PhD (13.4%). The lowest group was foundation (2.8%).

### Table 4. Education Levels of Respondents

| Education level | Frequency | Percent | Cumulative Percent |
|-----------------|-----------|---------|--------------------|
| Foundation      | 9         | 2.8     | 2.8                |
| Bachelor Degree | 163       | 50.9    | 53.8               |
| Master Degree   | 105       | 32.8    | 86.6               |
| PhD             | 43        | 13.4    | 100.0              |
| Total           | 320       | 100.0   |                     |

4.2 Measurement Model/ Confirmatory Factor Analysis (CFA)

First of all and before doing the Confirmatory Factor Analysis, the exploratory factor analysis (EFA) outcome should be tested for further analysis. The EFA provided six dimensions namely: fulfillment, interactivity, information, system availability, reliability, and privacy / security. To make statements coding and to test the reliability of these dimensions for further analysis, Cronbach alpha and KMO Bartlett’s test of sphericity were conducted as shown in Tables 5, 6, and 7.

### Table 5. Coding of Dimensions Statements

| Dimensions        | Statements Coding |
|-------------------|-------------------|
| Fulfillment       | FU1, FU2, FU3, FU4, FU5, FU6, FU7, FU8, FU9 |
| Interactivity     | INT1, INT2, INT3, INT4, INT5, INT6, INT7, INT8, INT9, INT10 |
| Information       | INF1, INF2, INF3, INF4 |
| System Availability| SA1, SA2, SA3, SA4 |
| Reliability       | RE1, RE2, RE3 |
| Privacy / Security| PS1, PS2, PS3 |
| Expected          | EXP1, EXP2, EXP3, EXP4, EXP5, EXP6 |
| Perceived         | PER1, PER2, PER3, PER4, PER5, PER6 |
| Disconfirmation   | DIS1, DIS2, DIS3, DIS4, DIS5, DIS6 |

### Table 6. Cronbach Alpha Reliability Test for CFA

| Dimension | No. of items | Cronbach’s alpha |
|-----------|--------------|------------------|

As shown in Table 6, Cronbach alpha values of all dimensions ranged from 0.790 to 0.889, which indicates that all dimensions were reliable. The minimum standard for Cronbach’s alpha is 0.7 [10, 11], which gives good internal consistency estimates. Table 7 indicates that KMO and Bartlett's Test are within the acceptable limits, where KMO is more than 0.5 and Bartlett's Test is less than 0.05 [12, 13]. Generally, the Cronbach alpha and KMO Bartlett’s test of sphericity confirm the dimensions reliability and their validity for Confirmatory Factor Analysis.

The normality assessment has been done by assessing the measures of skewness for every item. The absolute value of skewness 1.0 or lower indicates the data is normally distributed [14, 15]. However, bootstrapping data has been done to reach the confidence intervals, as well as the significance for every parameter involved in the analysis. After performing bootstrap 2000 samples at 95% confidence level, the regression standard error was replaced with bootstrap standard error and results were generated. The subsequent reported results are based on bootstrap standard error.

### 4.3 Measurement Model Fit

The measurement model plays an important role in dealing with the observed inputs and the observed outputs. The research model includes ten constructs with 57 items. The constructs comprise the six dimensions outcome of the EFA besides the expected, perceived, disconfirmation and satisfaction constructs. The (Confirmatory Factor Analysis) CFA for the model was conducted using AMOS21, and the measurement model was constructed as shown in Figure 6 and the fit indexes shown in Table 8.

| Table 7. KMO and Bartlett's Test for All Dimensions |
|-----------------------------------------------------|
| Kaiser-Meyer-Olkin Measure of Sampling Adequacy. 0.932 |
| Approx. Chi-Square 5801.421 |
| Bartlett's Test of Sphericity df 528 |
| Bartlett's Test of Sphericity Sig. 0.000 |

As shown in Table 6, Cronbach alpha values of all dimensions ranged from 0.790 to 0.889, which indicates that all dimensions were reliable. The minimum standard for Cronbach’s alpha is 0.7 [10, 11], which gives good internal consistency estimates. Table 7 indicates that KMO and Bartlett's Test are within the acceptable limits, where KMO is more than 0.5 and Bartlett's Test is less than 0.05 [12, 13]. Generally, the Cronbach alpha and KMO Bartlett’s test of sphericity confirm the dimensions reliability and their validity for Confirmatory Factor Analysis.

The normality assessment has been done by assessing the measures of skewness for every item. The absolute value of skewness 1.0 or lower indicates the data is normally distributed [14, 15]. However, bootstrapping data has been done to reach the confidence intervals, as well as the significance for every parameter involved in the analysis. After performing bootstrap 2000 samples at 95% confidence level, the regression standard error was replaced with bootstrap standard error and results were generated. The subsequent reported results are based on bootstrap standard error.

| Table 8. Statistics for Goodness of Fit Indexes For All Items |
|-------------------------------------------------------------|
| Absolute Fit Measures | Incremental Fit Measures |
|------------------------|-------------------------|
| P-Value | RMSEA | CFI | TLI | ChiSq/df |


The goodness of fit of the measurement model was measured to describe how well the model fits a set of observations. Goodness- of-fit index is a numerical summary of the discrepancy between the observed values and the values expected under the research model [16]. Statistics for goodness of fit indexes includes a minimal value for the discrepancy (CMIN), degree of freedom (DF), goodness of fit (GFI), root mean square error approximation (RAMSEA), adjusted goodness of fit (AGFI), normed fit index (NFI) and comparative fit index (CFI). Table 8, shows the statistics for goodness of fit indexes. As shown from the table, the fitness indexes did not meet the required factor loading, which should be 0.6 and above [14, 15]. There are several items have factor loading below 0.6 such as FU7, FU8, FU9, INT6, INT7, INT8, INT9, INT10, EXP5, US3 and US5, therefore, these items were deleted one by one with developing the measurement model until the fitness indexes were achieved. The sequence of item deletion is shown in Table 9. The sequence of deletion the low factor loading items led to develop the final measurement model as shown in Figure 7. The fitness indexes were examined to meet the required factor loading. The statistics for Goodness of Fit Indexes are shown in Table 10.

Table 9. The sequence of item deletion

| Absolute Fit Measures | Incremental Fit Measures |
|-----------------------|--------------------------|
|                       | P-Value | RMSEA | CFI | TLI | ChiSq/df |
| Acceptable Fit        | >0.05   | <0.05 | ≥0.90 | ≥0.90 | <5     |
| Obtained Fit For All  | 0.000   | 0.070 | 0.773 | 0.757 | 2.570  |
| Obtained Fit Deleted  | 0.000   | 0.069 | 0.784 | 0.769 | 2.529  |
| FU8                   | 0.000   | 0.068 | 0.798 | 0.783 | 2.458  |
| Obtained Fit Deleted  | 0.000   | 0.069 | 0.784 | 0.769 | 2.529  |
Table 10. Statistics for Goodness of Fit Indexes Final

| Incremental Fit Measures | Absolute Fit Measures |
|--------------------------|-----------------------|
|                          | P-Value | RMSEA | CFI | TLI | ChiSq/df |
| Acceptable Fit           | >0.05   | <0.05 | ≥0.90 | ≥0.90 | <5 |
| Obtained Fit             | 0.000   | 0.046 | 0.923 | 0.916 | 1.687 |

Figure 7. Final Measurement Model of The Relation Between e-SQ and User Satisfaction in Malaysian IHL

As shown in Table 9, all the obtained indexes values for Goodness of Fit are within the acceptable
values, which indicate that the measurement model fits closely the data and the factor loading of all items is above 0.6. The final measurement model shown in Figure 7 contains the ten constructs but with 46 items after deleting the items with low factor loading. The measurement model test confirms the adequacy of the measurement model, which should be measured for validity to develop the structural model. Construct validity is usually used for validity measurement.

4.4 Construct Validity

The construct validity refers to how well a test or tool can measure the construct that it was designed to measure [17]. CFA can be used to evaluate a model reliability and validity. One means of evaluating construct validity using CFA is to merge several constructs into a single model and assess the covariance pattern among factors representing the constructs against a pattern predicted from theory or basic knowledge about the relations among the constructs [18]. Confirmatory Factor Analysis (CFA) is required to confirm the underlying relationships among the latent variables, which cannot be observed directly [19].

This research utilized CFA on the model of six variables (six dimensions with their statements) generated by EFA, which are fulfillment, interactivity, information, system availability, reliability and privacy/security. All these dimensions with their statements were demonstrated previously [9]. The research aims to investigate the extent model constructs (dimensions) that can measure user satisfaction. Construct validity elucidates which items are convenient for each dimension, and confirms the dimension of the domain, which was operationally defined. The construct validity includes convergent validity and discriminant validity.

4.5 Convergent Validity

The convergent validity is the correlations between measures of the same constructs that are obtained with different measurement methods [20]. Convergent validity tests that constructs that are expected to be related are really related. The convergent validity is attained if there is a high correlation between items measuring the same construct, and the weak correlation confirms the likelihood that the items do not capture the same construct [21]. The convergent validity can be measured by several methods such as Cronbach alpha, composite reliability (CR) and average variance extracted (AVE) [22]. In this research, AVE and CR are used to measure the convergent validity. AVE is the average amount of variance in observed variables that a latent construct is able to explain [23]. AVE measures the variance that a construct has from its other measuring items comparing to the amount due to measurement error. AVE should be equal to 0.50 and above [22, 24, 25]. If the average variance extracted is less than 0.50, then the variance due to measurement error is greater than the variance due to the construct. In this case, the convergent validity of the construct is questionable. CR should exceed 0.7 [22]. As shown in Table 11.

| CONSTRUCTS         | Composite Reliability (CR) | AVE   |
|--------------------|----------------------------|-------|
| Expected           | 0.834                      | 0.504 |
| Fulfillment        | 0.875                      | 0.540 |
| Interactivity      | 0.832                      | 0.500 |
| System Availability| 0.798                      | 0.500 |
| Privacy/ Security  | 0.821                      | 0.605 |
| Perceived          | 0.883                      | 0.557 |
| Satisfaction       | 0.888                      | 0.666 |
| Information        | 0.856                      | 0.599 |
| Reliability        | 0.863                      | 0.677 |
| Expected           | 0.834                      | 0.504 |
Table 11 indicates that all constructs have values of AVE ≥ 5 and composite reliability of all constructs is also higher than 0.70, which fit the acceptable values and prove that the convergent validity is attained through the correlations between the measures of each construct.

### 4.6 Discriminant Validity

Discriminant validity is the extent to which latent variable discriminates from other latent variables [23]. Discriminant validity (or divergent validity) tests of constructs that should not have relationships, have not any relationship. Generally, the discriminant validity means that the measuring items of a construct measure what they expected to measure. This means the items should correlate higher among them than they correlate with other items from other constructs that are theoretically supposed not to correlate [25]. The discriminant validity is attained when there is no correlation between two different dimensions. The discriminant validity can be checked at item and construct levels. Chin criterion is used to check it at item level, while Fornell and Larcker criterion is used at construct level. The discriminant validity is also estimated by AVE-SV test, which introduced by Fornell and Larcker in 1981. This test compares the average variance extracted value (AVE) and the share variance (SV) (i.e., squared correlations) between dimensions [26, 27]. High discriminant validity is achieved when the AVE is greater than the squared correlation [17, 24]. The discriminant validity of all constructs is shown in Table 12.

In Table 12, the off diagonal values represent squared correlation and diagonal values represent the construct AVE [25]. The AVE values should be compared with the correlation coefficients of each construct and with the other constructs. The AVE values are greater than inter construct correlation for all constructs.

|                | Expected | FUL | INT | SA | PS | Perceived | Satisfaction | INFO | RE | DIS |
|----------------|----------|-----|-----|----|----|------------|--------------|------|----|-----|
| Expected       | 0.710*   |     |     |    |    |            |              |      |    |     |
| FUL            | 0.024    | 0.735*|     |    |    |            |              |      |    |     |
| INT            | 0.002    | 0.576 | 0.707*|    |    |            |              |      |    |     |
| SA             | 0.700    | 0.008 | 0.055 | 0.707*|    |            |              |      |    |     |
| PS             | 0.560    | -0.043 | -0.004 | 0.506 | 0.778*|            |              |      |    |     |
| Perceived      | 0.707    | -0.008 | 0.022 | 0.706 | 0.752 | 0.746*     |              |      |    |     |
| Satisfaction   | 0.702    | -0.058 | -0.012 | 0.704 | 0.587 | 0.701 | 0.816*   |      |    |     |
| INFO           | 0.703    | -0.016 | -0.020 | 0.688 | 0.575 | 0.731 | 0.732 | 0.774*|    |     |
| RE             | 0.018    | 0.733 | 0.512 | -0.054 | -0.047 | -0.011 | -0.042 | -0.082 | 0.823*|    |
| DIS            | 0.701    | -0.029 | 0.025 | 0.703 | 0.605 | 0.722 | 0.739 | 0.717 | -0.045 | 0.745*|

Note: * = Square root of AVE

### 5. STRUCTURAL EQAUTION MODEL (SEM)

Structural model is a model having paths reflecting causal dependencies between intrinsic variables, which can exclude the measurement model and linear regression [28]. The measurement model should be measured for validity to come out with structural model. Construct validity is usually used for validity measurement. The structural model was developed as shown in Figure 8. The goodness of fit index of the structural model was measured to judge the overall model and detect its performance. The results are shown in Table 13.
Note: FUL = Fulfillment; INT = Interactivity; INFO = Information; PS = Privacy/Security; SA = System Availability; RE = Reliability; Perceived = Perceived Performance; Expected = Expected Outcome; Satisfaction = User Satisfaction; and DIS = Disconfirmation

**Figure 8.** The Structural Equation Model (SEM) for e-SQ and Users Satisfaction Relationship in Malaysian IHL

**Table 13.** Structural Model Fit Measure Assessment

| Absolute Fit Measures | Incremental Fit Measures |
|-----------------------|--------------------------|
| P-Value | RMSEA | CFI | TLI | ChiSQq/df |
| Acceptable Fit | >0.05 | <0.05 | ≥0.90 | ≤0.90 | <5 |
| Obtained Fit | 0.000 | 0.047 | 0.9210.914 | 1.697 |

As shown in Table 13, all the obtained indexes values for goodness of fit are within the acceptable values, which indicate that the structural model fits closely the data, which confirms the structural model adequacy.

6. **HYPOTHESIS TESTING**

Hypothesis testing is the method of testing whether a hypothesis is likely to be true. The hypothesized relations between constructs were put in null hypotheses, which presume true. The P-value is used as a tool to evaluate how well the sample data support that the null hypothesis is true [29, 30]. It measures how research data is compatible with the null hypothesis. High P values indicate that research data are
likely with a true null, while Low P values indicate that research data are unlikely with a true null. Table 13 shows the regression weights of the constructs and the hypotheses. Based on this, Table 14 is developed showing clearly the hypotheses testing identifying whether each hypothesis is supported or not.

Table 14. Regression Weights: (Group number 1 - Default model)

| Construct | Path | Code | Estimate | S.E. | C.R. | P    | Remarks |
|-----------|------|------|----------|------|------|------|---------|
| Expected  | <--- | FUL  | -0.052   | 0.075| - 0.693 | 0.488| Not S   |
| Perceived | <--- | FUL  | -0.075   | 0.062| - 1.203 | 0.229| Not S   |
| Perceived | <--- | INFO | 0.316    | 0.062| 5.104  | 0.001| S       |
| Expected  | <--- | INFO | 0.43     | 0.077| 5.603  | ***  | S       |
| Perceived | <--- | PS   | 0.365    | 0.055| 6.587  | ***  | S       |
| Expected  | <--- | PS   | 0.104    | 0.06 | 1.741  | 0.082| Not S   |
| Expected  | <--- | SA   | 0.378    | 0.072| 5.238  | ***  | S       |
| Perceived | <--- | SA   | 0.35     | 0.06 | 5.82   | ***  | S       |
| Perceived | <--- | RE   | 0.099    | 0.058| 1.711  | 0.048| S       |
| Expected  | <--- | RE   | 0.133    | 0.07 | 1.888  | 0.049| S       |
| Expected  | <--- | INT  | -0.05    | 0.056| - 0.9  | 0.368| Not S   |
| Perceived | <--- | INT  | -0.002   | 0.046| - 0.045 | 0.964| Not S   |
| DIS       | <--- | Expected | 0.435 | 0.081| 5.371  | ***  | S       |
| DIS       | <--- | Perceived | 0.522 | 0.081| 6.428  | ***  | S       |
| Satisfaction | <--- | Expected | 0.303 | 0.103| 2.946  | 0.003| S       |
| Satisfaction | <--- | Perceived | 0.482 | 0.107| 4.518  | ***  | S       |
| Satisfaction | <--- | DIS   | 0.259    | 0.114| 2.262  | 0.024| S       |

Table 15. Hypothesis Testing

| Construct          | Hypothesis | Relations    | P-value | Remarks |
|--------------------|------------|--------------|---------|---------|
| Reliability        | H5B        | Perceived <--- RE | 0.048  | Supported |
|                    | H5A        | Expected <--- RE | 0.049  | Supported |
|                    | H4B        | Perceived <--- FUL | 0.488  | Not Supported |
| Fulfilment         | H4A        | Expected <--- FUL | 0.229  | Not supported |
|                    | H8B        | Perceived <--- INFO | 0.001  | Supported |
| Information        | H8A        | Expected <--- INFO | 0.001  | Supported |
| Privacy/Security   | H3B        | Perceived <--- PS | 0.001  | Supported |
|                    | H3A        | Expected <--- PS | 0.082  | Not supported |
| Interactivity      | H7B        | Perceived <--- INT | 0.964  | Not Supported |
|                    | H7A        | Expected <--- INT | 0.368  | Not supported |
As shown in Table 15, according to the corresponding high P-value, five hypotheses from all the constructs are not supported because the P-value is greater than 0.005 according to the Fit Measure Assessment in Table 13. Fulfilment and Interactivity constructs have not supported hypotheses. Also Privacy/Security construct has one hypothesis not supported. In summary, H4A, H4B, H3A, H7A, and H7B are not supported by the research data, thus they rejected.

7. HYPOTHESES TESTING EXPLANATION

As shown in Figure 9, the fulfillment dimension has a low effect on perceived performance, where the corresponding P-value is 0.488 greater than the acceptable Fit 0.005. The dimension is also has a low effect on expected outcome for the same reason (P-value is 0.229 greater than 0.005). Therefore, the fulfillment dimension is considered not significant. The dimension interactivity is also not significant due its minor effect on perceived performance and expected outcome, where the P-value is equal 0.964 and 0.368 respectively greater than the acceptable Fit. As a researcher’s idea; not all the dimensions should have an impact on the perceived and expected constructs. This reveals that the fulfillment and interactivity dimensions are not necessarily can formulate or affect the user satisfaction even they are significant factors for assessing the service quality. Service quality has strong impacts on customer satisfaction [31, 32], however, there is not a set of accepted factors capable to explain what satisfies the customer, because customer satisfactions vary according to culture and industry [33].

H2A: There is a positive relationship between system availability and expected outcome.

The P-value for the relation between system availability and expected outcome is 0.001, which is less than the acceptable Fit 0.005. Thus, this relation path is significant and supports the proposed hypothesis H2A. This indicates that the availability of the services can confirm the user expectations about the good quality of the provided services.

H2B: There is a positive relationship between system availability and perceived performance.

The P-value for the relation between system availability and perceived performance is 0.001, which is less than the acceptable Fit 0.005. Thus, this relation path is significant and supports the proposed hypothesis H2B. It can be concluded that good system availability can increase the user perception about the website services.

H3A: There is a positive relationship between privacy/security and expected outcome.

This relation is not significant because the P-value is 0.082, which is greater than the acceptable Fit 0.005. Therefore, this relation does not support the hypothesis H3A. This means that privacy/security has no effect on the user expectations because the user may accept the current security solution since
he/she cannot expect better solutions than the website.

H3B: There is a positive relationship between privacy/security and perceived performance.

This relation is significant because the P-value is 0.001, which is less than the acceptable Fit 0.005. Therefore, this relation supports the hypothesis H3B. This means that privacy/security affects the user perceived performance because the current security solutions provide good protection for his personal information and financial issues.

H4A: There is a positive relationship between fulfillment and expected outcome.
H4B: There is a positive relationship between fulfillment and perceived performance.

The relations between fulfillment and expected outcome and perceived performance are not significant because the fulfillment dimension is ignored as explained previously in this section. Therefore the relations do not support the hypothesis H4B and H4A. However, culture and industry play an important role to define the factors that affect user satisfaction [33]. Fulfillment dimension is a significant dimension in assessing e-SQ, however, it might lead to user satisfaction or dissatisfaction [34]. This is confirmed by this research where it proved it is significance in evaluating e-SQ and did not lead to user satisfaction based on hypotheses testing results.

H5A: There is a positive relationship between reliability and expected outcome.

The P-value for this relation is 0.049, which is less than the acceptable Fit 0.005. Thus, this relation path is significant and supports the proposed hypothesis H5A. This indicates that the increased website reliability offers an opportunity for the user to expect better quality of the provided services.

H5B: There is a positive relationship between reliability and perceived performance.

The P-value for this relation is 0.048, which is less than the acceptable Fit 0.005. Thus, this relation path is significant and supports the proposed hypothesis H5B. This reveals that the increased website reliability enhances the user perceptions on the quality of the provided services.

H7A: There is a positive relationship between interactivity and expected outcome.
H7B: There is a positive relationship between interactivity and perceived performance.

The relations between interactivity and expected outcome and perceived performance are not significant because the interactivity dimension is ignored as explained previously in this section. Therefore, the relations do not support the hypotheses H7B and H7A. However, it is difficult to determine the factors that affect user satisfaction. The insignificance relation may be due to lack of real interaction between IHL websites and users.

H8A: There is a positive relationship between information and expected outcome.

This relation is significant because the P-value is 0.001, which is less than the acceptable Fit 0.005. Therefore, this relation supports the hypothesis H8A. This indicates that little and blurry information can decrease users’ expectations on the services quality and may lead to reducing his interaction with the website.

H8B: There is a positive relationship between information and perceived performance.

This relation is significant because the P-value is 0.001, which is less than the acceptable Fit 0.005.
Therefore, this relation supports the hypothesis H8B. This indicates that simple and detailed information can enhance users’ perceptions on the services quality and increase the interaction with the website.

H10: There is a negative relationship between expected outcome and disconfirmation.

The p-value for the relation between disconfirmation and expected outcome is 0.001, which is less than the acceptable p-value 0.005. Thus, this relation is significant and supports the proposed hypothesis H10. This means that better user expectations can decrease the disconfirmation.

H11: There is a positive relationship between perceived performance and disconfirmation.

The p-value for the relation between disconfirmation and perceived performance is 0.001, which is less than the acceptable p-value 0.005. Thus, this relation is significant and supports the proposed hypothesis

H11. Disconfirmation is considered as a significant predictor of perceived outcome.

H13: There is a positive relationship between expected outcome and user satisfaction.

The p-value for the relation between expected outcome and user satisfaction is 0.003, which is less than the acceptable p-value 0.005. Thus, this relation is significant and supports the proposed hypothesis H13. It can be concluded that user expectations can contribute to formulate user satisfaction.

H14: There is a positive relationship between perceived performance and user satisfaction.

The p-value for the relation between perceived performance and user satisfaction is 0.001, which is less than the acceptable p-value 0.005. Thus, this relation is significant and supports the proposed hypothesis H14. This also indicates that user perceptions can contribute to formulate user satisfaction.

H12: There is a positive relationship between disconfirmation and user satisfaction.

The p-value for the relation between disconfirmation and user satisfaction is 0.024, which is less than the acceptable p-value 0.005. Thus, this relation is significant and supports the proposed hypothesis

H12. This also indicates that user disconfirmation can contribute positively on user satisfaction.

All the non-significant relations and dimensions will be deleted to come out with the final model. The value of R² (squared multiple correlations) for the model is the most important output in the Standardized Regression Weight [14, 15]. This means that R² is an indication of construct significance. R² for each item should have value greater than 0.4, otherwise it will be deleted [14, 15]. As shown in Figure 9, the values of R² for the main constructs are significant: for perceived performance is 0.85, for expected outcome is 0.73, for disconfirmation is 0.77 and for the satisfaction is 0.72. All these values considered as significant.
Note: NS = Not Significant; * = Significant; = Significant; > = Not Significant

**Figure 9.** Conceptual Model of the Relationship between e-SQ and User Satisfaction in Malaysian IHL (All Dimensions)

8. **FINAL E-SQ MODEL**

The final model of the relation between e-SQ and user satisfaction based on expectancy disconformation theory and after revising the hypotheses and the dimensions is shown in Figure 10.
As shown in Figure 10, the final model of the e-SQ and user satisfaction relation includes eight constructs. Four of these constructs are constructs (dimensions) of e-SQ, namely information, reliability, privacy/security, and system availability. The other four constructs are related to the disconfirmation theory. The development of the SEM led to the deletion of several items based on the reliability factors. It also led to the deletion of several hypotheses and two constructs namely fulfillment and interactivity based on the P-values. The final valid relations are illustrated in the final model in Figure 10.

The model indicates that the four e-SQ dimensions can affect the perceived performance, while three of them namely information, reliability, and system availability can affect the expected outcome except privacy/security. This means that user expectations do not affect privacy/security because user may trust the website as it is professional and has more experience in privacy/security solutions than the user. The model also indicates that perceived performance and expected outcome can lead to disconfirmation or satisfaction.

Generally, this model confirms the expectancy disconfirmation theory role in determining the user satisfaction. Consequently, this model can be used as a measure for user satisfaction of the e-service quality provided by the IHL in Malaysia.

9. DISCUSSION ON CFA AND SEM

The SEM is a statistical method designed to test a conceptual or theoretical model. This method provides a more thorough understanding on the study variables. A key feature of SEM is its ability to link observed indicators to latent variables [30]. The SEM includes two models namely the measurement model or confirmatory factor analysis (CFA), and the structural equation model. Confirmatory Factor Analysis (CFA) is one of the main factor analysis techniques. CFA attempts to confirm hypotheses and uses path analysis diagrams to represent variables and factors [28]. CFA can be used to evaluate a model reliability and validity and it is required to confirm the underlying relationships among the latent variables, which cannot be observed directly [19]. Confirmatory factor analysis (CFA) is used to validate the factorial validity of the models derived from the results of EFA [28].

The data required for SEM were collected by a survey questionnaire conducted with 320 students.
in the selected IHL. The questionnaire includes 10 constructs (dimensions) and 57 items. These constructs include the six dimensions and their 33 items extracted by the EFA, in addition to expected, perceived, disconfirmation and satisfaction constructs. The CFA was conducted on the selected construct and items using AMOS. Statistics for Goodness of Fit Indexes and factor loading were used as criteria to measure the goodness of fit of the measurement model. A factor loading for a variable is a measure of how much the variable contributes to the factor; thus, high factor loading scores indicate that the factors are better accounted for by the variables [28]. These criteria led to delete several items reaching the final measurement model. This model was measured for validity by using construct validity to come out with constructs, items and hypothesis that constitute the final structural model for measuring user satisfaction. The structural model goodness of fit index was measured to judge the overall model and detect its performance. To test whether a hypothesis is likely to be true, hypothesis testing method was conducted using P-value as a criterion and led to deletion five hypotheses and confirm the rest hypotheses.

10. DISCUSSION ON THE RELATIONSHIP BETWEEN E-SQ AND USER SATISFACTION
Customer satisfaction is significant for organizations because it can be used as a metric to manage and improve their businesses. It can affect mainly marketing orientations of organizations because it can decide user intentions and loyalty. However, it is difficult to measure and achieve it completely because it is imperceptible thing. Several studies confirmed the relationship between service quality and satisfaction. However, they did not agree on specific factors capable to achieve the user satisfaction.

The expectancy disconfirmation theory (EDT) is used widely to measure user satisfaction [7], based on the difference between user expectation and experience in perceived products or services [6]. Normally, expectations and perceived performance of e-SQ are substantial indicators of disconfirmation, which is a major indicator of satisfaction [7]. In any case, some service quality dimensions may not be substantial for user satisfaction, but they can essentially prompt dissatisfaction when they are performed inadequately [34]. Moreover, not all service quality dimensions are necessarily able to achieve user satisfaction. This issue is confirmed by this research, where fulfillment and interactivity dimensions were significant for e-service quality but have no effect on user satisfaction. Generally, the hypothesized conceptual model confirms that there is a strong relation between e-service quality and user satisfaction. However, the dimensions of e-service quality are different based on the user perception and expectation and based on the study filed.

11. CONCLUSIONS AND FUTURE WORK
Regarding higher education, the literature includes a little researches focusing on minimal number of e-services rather than an entire view of e-SQ. This research tried to fill this gap and develop a measurement model for the relation between e-service quality and users satisfaction based on the expectancy disconfirmation theory regarding users’ perceptions and expectations. A quantitative approach was utilized to collect the data and a conformity factor analysis for the proposed model has been done through AMOS 21. The results showed that the model fit indexes are in acceptable values and ready to do the Structural Equation Model (SEM). The final model will lead to achieve the users’ satisfaction, consequently, the competitive advantage of the universities in their

REFERENCES
[1] A. Shekarchizadeh, A. Rasli, and H. Hon-Tat, "SERVQUAL in Malaysian universities: perspectives of international students," Business Process Management Journal, vol. 17, pp. 67-81, 2011.
[2] I. T. I. Al-Nuaimi, A. K. B. Mahmood, L. T. Jung, and H. H. Jebur, "A review of e-service quality dimensions in user satisfaction,” in 2013 International Conference on Research and Innovation in Information Systems (ICRIIS), 2013, pp. 186-191.
[3] R. J. Angell, T. W. Heffernan, and P. Megicks, “Service quality in postgraduate education,”
Quality Assurance in Education, vol. 16, pp. 236-254, 2008.

[4] R. Ladhari, "A review of twenty years of SERVQUAL research," International Journal of Quality and Service Sciences, vol. 1, pp. 172-198, 2009.

[5] M. Suuroja, "Service quality—Main conceptualizations and critique," University of Tartu Economics and Business Working Paper, 2003.

[6] N. Elkhani and A. Bakri, "Review on “Expectancy Disconfirmation Theory” (EDT) Model in B2C E-Commerce," Journal of Information Systems Research and Innovation, vol. 2, pp. 95-102, 2012.

[7] S. Picazo-Vela, "The Effect of Online Reviews on Customer Satisfaction: An Expectation Disconfirmation Approach," AMCIS 2009 Doctoral Consortium, p. 6, 2009.

[8] I. T. I. Al-Nuaimi, A. K. B. Mahmood, E. E. Mustapha, and H. H. Jebur, "Development of Measurement Scale for Hypothesized Conceptual Model of E-service Quality and User Satisfaction Relationship," Research Journal of Applied Sciences, Engineering and Technology, vol. 12, pp. 870-884, 2016.

[9] J. F. Hair, W. C. Black, B. J. Babin, R. E. Anderson, and R. L. Tatham, Multivariate data analysis vol. 6: Pearson Prentice Hall Upper Saddle River, NJ, 2006.

[10] J. C. Nunnally, I. H. Bernstein, and J. M. t. Berge, Psychometric theory, 3rd ed ed.: New York: McGraw-Hill., 1994.

[11] J. F. Hair, R. E. Anderson, R. L. Tatham, and C. William, "Black (1998), Multivariate data analysis," ed: Upper Saddle River, NJ: Prentice Hall, 1998.

[12] A. Field, Discovering statistics using SPSS: Sage publications, 2009.

[13] Z. Awang, "A handbook on SEM: Structural equation modelling," Kelantan: Universiti Teknologi MARA, 2011.

[14] Z. Awang, A Handbook on SEM for Academicians and practitioners: MPWS Rich Resources., 2014.

[15] A. Maydeu-Olivares and C. Garcia-Forero, "Goodness-of-fit testing," International Encyclopedia of Education, vol. 7, pp. 190-196, 2010.

[16] C. M. Voorhees, M. K. Brady, R. Calantone, and E. Ramirez, "Discriminant validity testing in marketing: an analysis, causes for concern, and proposed remedies," Journal of the academy of Marketing Science, vol. 44, pp. 119-134, 2016.

[17] H. E. Tinsley and S. D. Brown, Handbook of applied multivariate statistics and mathematical modeling: Academic Press, 2000.

[18] B. M. Byrne, Structural equation modeling with AMOS: Basic concepts, applications, and programming: Routledge, 2013.

[19] T. Raykov, "Evaluation of convergent and discriminant validity with multitrait–multimethod correlations," British journal of mathematical and statistical psychology, vol. 64, pp. 38-52, 2011.

[20] K. D. Carlson and A. O. Herdman, "Understanding the impact of convergent validity on research results," Organizational Research Methods, vol. 15, pp. 17-32, 2012.

[21] n. d. Hec, "Methodology," ed: higher education commission, 2015.

[22] A. M. Farrell, "Insufficient discriminant validity: A comment on Bove, Pervan, Beatty, and Shiu (2009)," Journal of Business Research, vol. 63, pp. 324-327, 2010.

[23] Y. Liu, "Developing a scale to measure the interactivity of websites," Journal of advertising research, vol. 43, pp. 207-216, 2003.

[24] A. Zait and P. S. P. E. BERTEA, "Methods for testing discriminant validity," Management & Marketing Journal, vol. 9, 2011.

[25] D. Gefen and D. Straub, "A practical guide to factorial validity using PLS-Graph: Tutorial and annotated example," Communications of the Association for Information systems, vol. 16, p. 5.
[26] J. Henseler, C. M. Ringle, and M. Sarstedt, "A new criterion for assessing discriminant validity in variance-based structural equation modeling," *Journal of the academy of Marketing Science*, vol. 43, pp. 115-135, 2015.

[27] A. G. Yong and S. Pearce, "A beginner’s guide to factor analysis: Focusing on exploratory factor analysis," *Tutorials in Quantitative Methods for Psychology*, vol. 9, pp. 79-94, 2013.

[28] A. Field, "Discovering statistics using SPSS for Windows: Advanced techniques for beginners (Introducing Statistical Methods series)," ed: SAGE publications Ltd, 2000.

[29] J. Valluzzi, S. Larson, and G. Miller, "Indications and limitations of structural equation modeling in complex surveys: Implications for an application in the Medical Expenditure Panel Survey (MEPS)," in *Joint Statistical Meetings-Section on Survey Research Methods, Agency for Healthcare Research and Quality*, 2003, pp. 4345-4352.

[30] H. Li and R. Suomi, "Dimensions of e-service quality: an alternative model," in *Future Generation Communication and Networking Symposia, 2008. FGCNS’08. Second International Conference on*, 2008, pp. 29-35.

[31] X. Zhang and V. R. Prybutok, "A consumer perspective of e-service quality," *Engineering Management, IEEE Transactions on*, vol. 52, pp. 461-477, 2005.

[32] I. N. Ibok and A. S. John, "Investigating Customer Satisfaction Driven Values in the Retail Banking Industry," *International Journal of Finance and Accounting*, vol. 2, pp. 292-296, 2013.

[33] Z. Yang and X. Fang, "Online service quality dimensions and their relationships with satisfaction: A content analysis of customer reviews of securities brokerage services," *International Journal of Service Industry Management*, vol. 15, pp. 302-326, 2004.