International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2019): 7.583

Online Survey-based Research Using Likert-Scale Questionnaire as an Educational Tool during COVID-19: From Design and Development to Validation Process and Interpretation

Shahid Hassan

School of Medicine/IMU Centre for Education, International Medical University, Malaysia

Abstract: ***Background**: It is crucial for the faculty in teaching institution to research and publish during COVID pandemic besides, teaching and assessing in medical education. It is an opportunity for faculty to conduct questionnaire based research and learn about survey based research design and its validation. It has been found that most of the researchers rely on published questionnaire though; designing a questionnaire is a good research practice to start with. It may pay off the researchers in many ways in future. Keeping with researcher's incentives of readers and difficulties they may encounter in creating questionnaire-based research and its usability, author is obliged to write this paper and share his knowledge and skills with those interested to learn, how to produce validated instruments for survey-based research. **Objective**: To learn how to create a validated questionnaire from design to validation process and its usability in research. **Method**: Emphases on questionnaire-based research in this paper will focus on creation of items in a questionnaire using survey design and its validation using a fictitious data. In current questionnaire-based survey research data is collected on a fifteen items scale score measured on a Likert scale from 1 to 4, where 1 represents strongly disagree, 2 disagree, 3 agree and 4 strongly agree. Paper will discusses the entire steps of validation of questionnaire-based research and its usability in detail. **Result**: A flow chart developed in the process portrays the entire process of creation, validation and usability of questionnaire. A fictitious data created to write this paper has generated a number of tables including content validity indices, reliability indices and a number of factor analysis matrix. **Conclusion**: The current paper described how to create a questionnaire based survey-research and its validation in detail, which may help readers and new researchers to understand descriptive type of research using survey design, specially explored for faculty interested to develop their research skills during COVID-19 pandemic and share their experiences with global community.*

Keywords: Questionnaire-based research, Survey-design, Validation process, Reliability, Validity, Factor analysis, Usability

1. Introduction

In an unprecedented challenging situation of COVID-19 pandemic that we all are currently facing in delivery of education globally, teaching faculty has the additional responsibilities of researching and publishing besides, teaching and assessing in medical education[1]. Multiple subject areas are opened for research in teaching and learning and often new researchers are inclined to perform questionnaire based research using the inventories borrowed from published literature. However, researchers find difficulties in establishing suitability of those inventories in a changing socio-cultural environment and other associated issues. The problem further mounts for researchers when those established in literature questionnaires are to be modified to determine the variability of the scale inventory for a targeted population. Instead of research conducted using established questionnaire that often has the issue of granting permission from the authors or journal administration, a teacher might research to establish his own validated survey instruments. Keeping with researcher's incentives, knowledge and difficulties encountered to create questionnaires and conduct research using questionnaires, author is obliged to write this paper and share his knowledge and skills with global community in learning how to produce validated instruments for survey-based research.

2. Method

Validation of a questionnaire created for survey-based research involves multiple steps from its inception of creative idea to design, development and validation of questionnaire. Validation process of a questionnaire is a standard procedure (see figure 1) that must be followed through in a questionnaire accepted for valid research and data collection. This model is used here to describe each aspects of validation in a friendly user manner to encourage researchers to take this initiative of producing validated instrument for research and sharing of knowledge during the COVID pandemic. Authors got involved into this situation after a number of consultancy sessions with faculty seeking review of their research proposals based on published questionnaire with a number of issues not realised by researchers. This may include authors or journal permission, suitability for targeted population, revalidation, analysis outcome and interpretation of data obtained from that questionnaire-based survey.

In survey based research data is collected using a questionnaire with multiple items measured either on a dichotomous option of, "Yes or No" or on a Likert scale from, "1 to 4", where 1 represents strongly disagree, 2 disagree, 3 agree and 4 represents strongly agreed (see appendix). The original Likert scale is a set of statements (items) offered for a real or hypothetical situation under study[2]. Participants surveyed are directed through an instruction to indicate their level of agreement (from strongly disagree to strongly agree) with every question (items) on a metric scale. The statements (items) in combination reveal the specific dimension of the construct represented by inter-linked with each other[3]. The issue here may arise, how to quantify these subjective preferential choices in a validated and reliable manner and that is what has been offered by a questionnaire based survey designed with a Likert scale[4, 5].

The process of creating a questionnaire using a Likert scale primarily will require validation in terms of content validity index (see figure 2) and a factor reduction technique in order to identify relevant items in relevant factors (construct) called factor analysis or more precisely Principal Component Analysis (PCA). This process analyses the data for initial extraction and subsequent rotation to identify factors with their respective items, which may not be the same a researcher might have created. To get started with selecting a rotation method one has to start with factor analysis procedure for our all the items in questionnaire designed, either as one factor or assigned to different factors with certain number of items in each factor. Current paper is an effort to provide an easy training to those interested in questionnaire based survey research. The process of creating and validating a newly created questionnaire are described as following.

2.1 Questionnaire Design

Let's begin with research design for its three basic flavours of being exploratory, descriptive and explanatory depending on intent of researcher. The difference in three types are:1. Where one aims to explore in an area that one knows nothing or have little knowledge.2. Where one wants to describe further with accuracy about something one already knows. 3. Where researcher wants to explain causal/predictive relationship among the variables. Under these three types of research we may have research design called experimental, survey, comparative, case study, observational, action research or mixed method research. Survey research falls primarily under descriptive research. The basic idea in this type research is to use or create a standardised instrument to collect data from a large number of respondents using interview-based, internet-based or questionnaire-based survey. Survey research is used for descriptive purposes although implication of this kind of research may be causal or predictive however, process of research is descriptive in nature.

After accomplishing the literature search relevant to subject area, researcher may start writing the items that covers every aspect in terms of content of the questionnaire for survey. First thing is to think of design of survey questionnaire based on the objective. If the objective is to gather the data for descriptive analysis, which basically requires to look at mean and standard deviation or percentage/proportion for data interpretation of items in different constructs depending on, whether questionnaire is one or multiple factors survey.

2.2 Questionnaire Validation

1) Data Collection and Cleaning

After the questionnaire is set and agreed upon among the members of research team (see appendix), next step is to run a pilot study among the subjects from the same environment that a larger survey is aimed at. Pilot study in survey research is crucial for the faculty in teaching institution to research and publish during COVID pandemic besides, teaching and assessing in medical education. Pilot study basically explores the clarity, language, connotation and understanding issues of each item by the respondent, who are also asked to provide comments if any difficulty found in response to each item in questionnaire. A follow through meeting of research team addresses those issues and fine tune the questionnaire ready to administer for validation purpose. Once the data is collected using online or hard copy response a work sheet is generated on Microsoft Excel called the, "Raw Data". Subsequent steps involved are as following.

2) Coding the Raw Data with Numeric Values in Excel:

As indicated in the Liker scale see what words have been assigned to each option from 1-4 (scale used in current study)and prepare to assign numeric values to each options written in words from strongly agreed to strongly disagreed using the steps as shown for recoding items with negative connotations, however, positively written questions is a better option. For coding the raw data to assign numbers following steps are followed through.

a) Hold the control key and click one column on the top of Excel sheet to highlight that column and while continuing to press control key, also press command key and click all the columns from Q1 -Q15.
b) Next press the central key and press letter "K" on key board, representing find and replace option.
c) In the table find the option of, "Find" and write the word exactly it is written in the Liker scale options of strongly agree to strongly disagree and press option of replace below.
d) This opens up option of replace with. Now enter the number you want to code and then click replace all and this will show all the cells changed to their numeric number as commanded and now press ok. This will show a popped up box indicating number of cells changed from text to their respective numeric numbers. Now repeat the same procedure for text in Likert scale until the entire raw data is changed to numbers.

3) Recoding the Raw Data with Negative Connotation in Excel:

Read through each items and identify items written with positive and negative connotation by denoting positively written with (N) and negatively written with (R). Now change those to a reverse order in Likert scale using the following steps.

a) Hold the control key and click one column at the top to highlight that column and while continuing to press control key, also press command key and click all those questions negatively connotated.
b) Next press the central key and press letter "K" on key board, representing find and replace option.
c) In the table find the option of find and there enter the word from exactly it is written in the Liker scale and press option replace below.

Paper ID: SR21512174451 DOI: 10.21275/SR21512174451

d) This opens up option of replace with. Here enter the number you want to be reversed coded then click replace all and this will show a popped up box indicating exactly the number of cells changed as commanded and now press ok. This will assign reverse coding to all the items as asked.

4) Cleaning the Data of Unengaged Respondents in Excel:

Cleaning the data from those respondents, who took the survey very casually are actually those uncommitted or unengaged respondents and should be removed from the data. Steps to identify unengaged respondents are as followed.

a) Choose a new column and give it a header of standard deviation (SD).
b) Select the cell below and type, = on key board and then type, "STD" and from the drop down menu select STDEV.S to open the parenthesis and highlight the 1st respondent entire row (all the items) and click enter.
c) Now fill in the rest of the cells in column (SD) by dragging the 1st cell down and look for standard deviation values with variations from one respondent to another respondent. Check out cells with 0 values and remove them from the data since these are the respondent who choose one option and filled out all the items with that option without applying their thoughts to logically decide, which applies the best.

5) Cleaning the Data for Missing Value in SPSS:

Reaching this stage in Microsoft Excel, it is time to transfer the data to SPPS to identify the missing value, which is a little difficult exercise in Excel to perform however; it is an easy process to do in SPSS. But before explaining how to find missing value in the data it will be better to describe steps to import the data file from Excel to SPSS. Collecting data initially in Excel is handier since it can directly download from online survey or can be typed down from the hard copy paper response by the participants. Subsequently it will be transferred from Excel to SPSS using the following steps.

a) Identify the location of Excel file from where it will be imported to SPSS. Prepare the Excel worksheet for import to SPSS by ensuring the rows contains the variables and column contains the subject. Also look for data not suitable for SPSS and remove them from the Excel sheet. Be informed that SPSS accepts number not text and therefore coding of Liker scale in Excel to numeric values using the method described above is done earlier.
b) In SPSS open file and go to, New and clickData. This opens a new SPSS data sheet.
c) Now in new SPSS data sheet go to file and from the dropdown menu go to import file (in new SPSS version)or open data (in old SPSS version) and click open file location as identified from source file of one's computer.
d) Having identified the Excel file highlight the file source and click open. This will open worksheet location option to choose and one may choose sheet 1 by default. Here check off the read variable names and percentage value at 95 and click open.
e) Excel file promptly will be imported to SPSS worksheet for further work to do the data cleaning, normality of data distribution, frequency output and the most wanted reliability indices as Cronbatch's alpha and correlation coefficient either Pearson's or Spearman's correlation coefficient depending on normality test of distribution whether met with or not respectively.

6) Identifying Missing Values in SPSS:

Step 1: Go to transform in menu bar and click recode into same variable.
Step 2: Transfer all the variables (items) into numeric variable box and click old and new variables box below.
Step 3: In old value box check off system missing
Step 4: In new value box enter the missing value (any number unlikely to be a data set value such as 999 and check off add, click continue and then ok.
Step 5: Go to missing column in variable view and hit the three dots to open the missing value drop box and check off discrete missing value and click ok.
Step 6: To allow same missing values to all the responses click missing value and copy it. Next highlight all the variables below and paste.

7) Looking for Normality of Data Distribution in SPSS

Before we embark on SPSS statistical test it is a good idea to go for and check the distribution of normality since this will be required as an important assumption to meet with as required for almost all the SPSS statistical methods (see figure 3). There will be many other assumptions associated with individual statistical test and those are discussed all along the test as we proceed in this write up. Following is the steps to perform normality of distribution test in SPSS.
Step 1: Click Analyse in SPSS drop down menu and go to descriptive statistics and click explore and this will open explore dialogue box.
Step 2: In explore dialogue box move the variable to be tested into the dependent list in explore dialogue box and if you have independent variable move it into factor box (optional).
Step 3: On the right side of dialogue box click statistics and leave confidence interval for mean at 95% by default and click continue.
Step 4: Click on plot to open explore plots dialogue box and here check off normality plot with tests and histogram. Now click continue followed by ok. Output table generated will be, test of normality showing Kolmogorov Smirnov and Shapiro-Wilk with significance value suggesting to accept or reject the null hypothesis for data being normally distributed depending on significant (<.05) or unsignificant (>.05) *p* value.
Step 5: For visual impact same can be observed for histogram (see figure 3) and QQ plot (data points with equal number above and below the line. There is another ways to establish the normality of distribution and that is to divide skewness/kurtosis statistics value by standard error value in descriptive table. A resulting value within -1.96 and +1.96 suggest data is normally distributed. For outlier do check the box plot and anything without an asterisk is considered ok.

8) Validity of Questionnaire

For validity of Likert scale questionnaire-base survey content validity index (CVI) seeks expert's opinion for basically relevancy of items to the subject area content and the specific construct created by the research team. CVI generates multiple indices including I-CVI (Individual-CVI) S-CVI (Scale CVI) with its further two indices called S-CVI/Ave (Scale CVI Average) and S-CVI/UA (Scale-CVI Universal Average (see figure 2)

Content Validity Index (CVI):
To measure the validity of a scale first thing is to define the construct and develop the items to measure the construct determined as the content validity. Content validity index is the degree to which an instrument has appropriate sample items for the construct being measured[6]. It determines, whether or not items sampled for inclusion on the tool adequately represent the domain of content addressed by the instrument[7] Content validity index is evaluated through individual (I-CVI) and scale related (S-CVI) after collection of experts response on excel format (see figure 2). Following measure in a systemic order is carried out as under.

I-CVI (Individual CVI)
I-CVI measure the efficacy of item as the proportion of experts giving an item a relevance rating of 3 or 4 in a 4-point Liker scale determined as the total agreement on item divided by the total number of experts (see table 1).

S-CVI (Sample CVI)
S-CVI measure the efficacy of the scale in toto as the content validity of overall sample determined by an average of I-CVI and this can further be divided into two (see figure 2)

a) **S-CVI/Ave (Average):**
   Average of the I-CVI for all items on the scale in a construct as one of the components of a questionnaire.
b) **S-CVI/UA (Universal Average):**
   Proportion of items on a scale that achieves a relevance rating of 3 or 4 by all the experts in a 4-points Likert scale. It's a conservative approach and is a very useful index.

Guidelines for Content Validity Index (CVI) Measure:
Most important informative procedure is to compute S-CVI for both of its types, S-CVI/Ave and S-CVI/UA. Excellent content validity is judged for I-CVI that meets the Lynn's (1986) criteria as following. There are no specific cut-off values established for CVI acceptable criteria however, Lynn's criteria provide a practical guideline to make logical decision as under.

Lynn's Criteria:
Excellent I-CVI with 3-5 experts = 1.00
Excellent I-CVI with 6-10 experts = 0.78 and above
Acceptable S-CVI with 6-10 experts = 0.90 and above

If Lynn's criteria are not met it indicates construct not adequately covered in the initial round of experts meeting and a second pool of experts' judgment becomes necessary

Steps to Calculate CVI:
A template for relevancy is prepared and sends out to selected experts with rubric (see figure 2). Ideally 10 or more experts should be selected for structure feedback using a template. However, 5-10 experts may be a good range to move on. For clarity simplicity of questionnaire items another template using dichotomous rubric of yes and no may suffice. Here more complex calculation of relevancy of items for I-CVI, S-CVI/Ave and S-CVI/UA will be described using the following steps.

1) First thing is to determine the number of agreement among the experts submitted as a documented form. At least 5-6 experts should be involved however, 10 experts would be ideal. Experts are provided with in-depth description of the construct to refer to and to make a logical decision on each item. This is calculated using following function in Excel. In case of a 4-point Likert scale a proportionate agreement will be 3 and 4 and in case of 5-point Likert scale it will be 4 and 5. Following function in Excel will determine the item number of agreement by typing.
   =count if (range of all experts rating, ">=3") and presenter. A numeric number will indicate the number of agreement among the expert on that items. Scrolling down from this number will fill in the remaining items cells on Excel sheet.
2) To calculate I-CVI the formula is, =number of agreement/number of ratters.
3) To calculate the S-CVI/Ave the formula is, = Average I-CVI range of all the items.
4) To calculate the S-CVI/UA, first we need to determine the total agreement among the expert from the list of agreement (using the example of 6 experts (see table 1) using the following function.
   =Count if (range of agreement, ">=6). The number achieved using this formula is divided by the number of ratters (Number of total agreement/ Number of total ratters)

This way we can determine the relevancy of items to its construct based on experts' opinion in terms of, I-CVI, 2. S-CVI/Ave and S-CVI/UA.

Clarity or simplicity of items can use yes or no criteria and will be sorted out addressing the comments associated with each items.

9) Reliability of Questionnaire

a) Cronbach's alpha:
Reliability is the extent to which a measurement is consistent if the test is repeated however, consistency to know about items in questionnaire is about the score across the items in the construct.Cronbac's Alpha is more appropriate for nominal interval data and it evaluate the extent to which different items on questionnaire measure the same ability or trait (see table 2). If items do not measure the same characteristics of a construct in a questionnaire, which may not be consistent internally. The coefficient of reliability with Cronbach's alpha is 0-1, negative value is theoretically not possible therefore, reported 0.

b) Reliability: Item-total correlation:
It refers to positive and strong relationship of items to item and items to total score of all items in a test (see table 3) and can be interpreted as:

- **Pearson's correlation coefficient (r)/Spearman's Correlation Coefficient (rho)**
  It is the correlation between item-to-item ranges from -1 to 1 and greater the number effective the item. Pearson's versus Spearman's Correlation Coefficient is decided depending on the normality of assumption achieved or not (see figure 3, table 4). If normality of distribution assumption is met with Pearson's correlation else, Spearman's correlation test is carried out.
- **Corrected item-total correlation**
  It is the correlation between each item and a scale overall score that excludes that item, ranges -1 to +1 and greater the number effective the item.
- **Cronbach's Alpha if item deleted**
  It is the value of overall alpha if the item is not included in the calculation and the value should be around the overall alpha (see table 2).

Steps to Perform Cronbach's alpha in SPSS:
Step 1: In drop down menu in SPSS on the top click Analyse, then scale and the Reliability Analysis.
Step 2: Transfer variables (items) q1, q2, q3.....qi and leave the model set as Alpha
Step 3: In the dialogue box click statistics.
Step 4: In the box description, select item, scale and scale if item deleted. In the interitem box, select correlation
Step 5: Click continue and then ok to generate the output.

To interpret the output, one can follow the rule of George and Mallery (2003)

No	Coefficient Alpha	Interpretation (Remarks)
1	>.9	Excellent
2	>.8	Good
	>.7	Acceptable
3	>.6	Questionable
4	>.5	Poor
5	<.5	Unacceptable

Steps to Perform Correlation Coefficient in SPSS:
Step 1: On the top of SPSS click Analyse and go to correlate and click bivariate.
Step 2: Move the variables to be tested for correlation to variable box on the right
Step 3: Check off Pearson or Spearman under the correlation coefficient and click ok. Correlation coefficients output table will generate for interpretation. A coefficient of .3 and more is considered good correlation between the two variable however, any value more than .8 indicate redundancy or multicollinearity and one of the two variables associated with very high correlation should be removed from the items list.

10) Factor Analysis of Questionnaire:
Factor analysis tells about how the items can be divided off through factor loading to determine different construct initially thought by the expert. There are two main approaches to factor analysis based on intent of a researcher, exploratory or confirmatory. Exploratory factor analysis intent to measure the dimensionality and often used in the early stages of research to establish the interrelationships among a set of items (variables) in a questionnaire[8]. Whereas, the confirmatory factor analysis is a set of techniques applied to test specific hypotheses or theories concerning the structure underlying a set of variables[9, 10] Factor analysis can be principal component analysis (PCA) or standard factor analysis (SFA). In PCA we work with original variables to produce to smaller set of variable with stronger linear correlation and this provides a practical model. In SFA we use a mathematical model of shared variance versus total variance and it provides a theoretical model. The two can be summarised in a way that if we are interested to develop a theory or looking for a theoretical solution and want it to be uncontaminated with variables, factor analysis is the best. If we are interested in real world scenario with empirical summary then principal component analysis is the right choice. In PCA outcome are interrelated factors called components and in factor analysis outcomes are truly factor as independent variables.

Factory analysis is based on the assumption that all variables (items) correlate to each other to some extent. Depending on research design variables should be measured at descriptive or inferential level of statistical test. The sample size for factor analysis is supposed to be big enough, over 200 but an acceptable range between the subjects and items should be in 10:1ratio[11]Factor analysis as a principle require big data and using smaller sample size results cannot be generalized. A big sample over 200 are basically required however, in terms of subjects and items ratio, 5:1 subjects/items are acceptable. In factor analysis it is important to know what we are measuring and how we are measuring. For how we are measuring there has to be in -depth criteria. In factor analysis criteria is provided by, how we explain the variances. There are three measures to determine our efficiency in factor analysis.

a) Kaiser's criteria, which uses eigen value and an eigen value of 1.0 is ideal which determines the amount of total variance explained by the tractor.
b) Scree test, which is the graphic representation of eigen value, shown in the shape of a curve that changes direction and becomes horizontal.
c) Parallel analysis, which is a quality control check and it compares the size of eigen value collected from our data with another randomly generated data with eigen value of same size. SPSS do not provide option for comparative analysis and can be downloaded as an application (Monte Carlo PCA for parallel analysis) on Microsoft Windows.

11) Steps to Perform Factor Analysis:
To start Factor analysis go to analyse in SPSS bar list and next go to dimension reduction and strike factor to open the dialogue. First move all the items intended for rotation into variable box. In order to perform rotation and define factors that we are interested in, we have to look at the other options in factor analysis first.
a) Under descriptive leave the initial solution checked off by default in statistics section. In correlation section we check off coefficient, determinant and KMO with Bartlett's test of sphericity in correlation matrix.
b) Under extraction we select Principles Component analysis (PCA) and check off Scree plot in extraction section. In analyse section correlation matrix and in display section unrotated factor solution are left checked off by default.

c) Under the score no option changes are made if we are not interested in saving the output table.
d) Under the choice of option we check off sorted by size in coefficient display format. This is an important step to know which item is loaded to which factor. In missing values exclude cases listwise is left checked off by default.
e) Lastly take a look on rotation and in rotation method select one of two options of Direct Oblimin in oblique rotation choice or Varimax in orthogonal rotation choice depending on whether data has the items correlated or uncorrelated. Uncorrelated method is easy to report as items are independent and easy to interpret.

12) Output Table in Factor Analysis

A researcher may start with larger number of items and using item analysis can reduce to lesser number in refining the questionnaire through dimension reduction technique of factor analysis so that the constructions become more meaningful. Once we click ok after every option and method in factor analysis is rightly checked off, it generates the output table and the first table to read is the correlation matrix.

In PCA items should have a correlation of 0.3 and above. Below 0.3 factor analysis may not be valid and another way to look at it is the determinant value, which has to be greater than 0.0001. At determinant value of less than 0.0001 items are considered not correlated. Another area of concern in correlation is when the items are highly correlated called multicollinearity and a cut off point for multicollinearity is a correlation value of more than 0.8 in a questionnaire based survey. In case of multicollinearity one of the items should be removed from the questionnaire. Another way to reconfirm low correlation items as measure of sample adequacy in factor analysis is to look at the Anti-image matrices. Here one looks at all the diagonal values (see table5) and a high diagonal value indicate that item can still be retained in questionnaire to continue with further steps in PCA. In current example items 3 and 6 have been observed to have very low correlation (light yellow colour in table 5) but their Anti-imagine value (dark yellow colour in table 5) have been quite high and can be retained in questionnaire.

More reliable way to look at the inter-item correlation is the KMO and Bartlett's test table (see table 4). A KMO value of >0.5 is acceptable for questionnaire to be factored. Similarly in Bartlett's test of sphericity we want a significant value of p = <0.001. The other assumption of data fitness for factor analysis is to look at the descriptive table generated with values of skewness kurtosis and dividing these values by standard error values. If result yields a score that fall within the range of -1.96 to +1.96, data is considered normally distributed to meet the assumption (see table 5). A communality table generated next gives the initial value of 1 to all the variables followed by value extracted and shared by each variable (item) as the index of communality (see table 5). After identifying data fitness using KMO and Bartlett's test of sphericity and normality of distribution measured through skewness and kurtosis in descriptive statistics next is the first extraction output of communalities. In factor analysis initial extraction in communalities table is established as 1 and the extracted value of each item against 1 correspond to R^2 and it determines, how much an item shares with rest of items to be important in factor analysis (see table6)

In extraction section of output table, Screeplot graphically tells us the number of factors extracted and some of those underlying potential factors (see figure 4). Same output is elaborated by total variance explained table in terms of factors explained (see table 7). Higher the number better it is.

Next is the output table is the component matrix both, unrotated and rotated. Other tables generated in this section may be pattern matrix and structure matrix. Pattern matrix are not much different from the component matrix however, structured matrix is more about correlation of factor among each other.

Once the number of components are determined it is time to interpret them and do that components are rotated determined by loading pattern in component matrix(see table 8) which tells us about variables clumped together in components or the factors referring to specific constructs named by the researchers (see table 9, before and after rotation). Factor analysis or the principle component analysis is all about rotation to achieve different factors or components and therefore a better understanding of rotation procedure is imperative to make sense of PCA.

13) Understanding Rotation Method in Factor Analysis:

The process of defining factors in the statistics is referred to rotation. Rotation in factor analysis is a mathematical procedure that rotates the factor axis in order to produce results that facilitates interpretation. Using rotation loading pattern becomes clear and easy and a more pronounced. In factor analysis rotation is associated with score scale and attempt is to reduce multiple items into clearly defined factors and this is accomplished through factor analysis. Following are the important steps in rotation methods in PCA.

a) Here we have several methods classified into orthogonal and oblique rotation. In orthogonal we have Quartimax, Varimax and Equamax. These options are used when we expect factors to be correlated and this can be checked by getting started with Direct Oblimin output correlation matrix (see table 4) and if a correlation of less than 0.32 is achieved we use one of those three, preferably the Varimax method. Other option is oblique method and it includes Direct Oblimin and Promax and we use these when factors are correlated. Using this process we define what rotation method has to be employed.
b) The whole purpose of a rotation is to create a simple structure, which helps in interpretation, which make sense of factor loading. However, the factor analysis will identify the factors with certain number of items in it but will not determine what construct it measure as this is beyond the pre-review of the factor analysis. The suitable names to factors are given by the researchers based on name reflecting the items under its fold (see table 9).
c) Next we select the rotation method knowing that first three methods are applied when factors are uncorrelated and last two methods are applied when factors ae

correlated and shows a correlation coefficient of more than .32.

d) To start, which orthogonal or oblique method to select, we first use Direct Oblimin method in oblique rotation assuming the factors are correlated. So we check off the direct Oblimin method and in display box we check off rotated solution and keep the maximum iteration up to 25 then click continue and press okay.
e) Using a Direct Oblimin method we directly move down to last component of correlation matrixin output table and look for correlation between factors whether greater or less than .32 analysed data (see table 4). Ignoring the diagonal values, which as a rule is always 1, we focus on values greater than .32 and having found one oblique method is continued. However for all practical purpose we may select one of three orthogonal methods, most commonly the Varimax leaving all the setting as it is and click continue and ok.
f) In output table, we move down to rotation component matrix and look for the simple structure. By definition a, “simple structure” is an item that has a significant loading and as many as zero loading as possible. A zero loading is defined as a factor loading which is greater than negative 0.1 and less than 0.1. There are many definitions of significant loading but a most practical definition of a significant loading is that a factor with factor loading of 0.3 (see table 7).However, some researchers may take significant loading value of 0.4 or 0.45.
g) Now having defined simple structure that has a significant loading and as many zero loading as possible besides, there may be many other values as well. A complex variables is defined as one with factor loading of 0.3 on two or more factors (see table 8).

Keeping the definitions of simple structure and complex variable, we interpret the factor analysis result and ultimately fine tune the questionnaire ready for research.

14) Questionnaire Factors Utility

It is also part of developing questionnaire to know how to use and analyse factors obtained from PCA. It depends on the objective of factor analysis and this can be univariate or multivariate. In univariate objective is to determine factors related to a construct. Univariate factor analysis does not mean that factor extracted is only one, rather it is about the factor not interested to establish relationship to other variables otherwise it will become a multivariate analysis. The objectives of univariate factor analysis therefore are,

a) To determine the factors representing a specific construct in a questionnaire. This usually will have the data collected on a Leikert scale and this may refer to three likely situation in exploratory factor analysis (EFA) or confirmatory factor analysis (CFA) as under,
   - Studies that do not find literature evidences or a theory available to explain and this becomes an explanatory factor analysis that seeks expert opinion.
   - The study that requires using a questionnaire from literature and factoring analysis is performed to know the social-cultural variation and adjustment of such questionnaire prior to performing PCA. This is important since that questionnaire may or may not be suitable for certain targeted for population in a confirmatory factor analysis.
   - Studies in which a questionnaire borrowed from the literature is modified for its content (items) to determine the variability of the scale inventory for a targeted population. Chi-squared test will be a good statistics to achieve the objective.
b) To determine the level of satisfaction obtained from the factor analysis either on awareness or perceptions of respondent. Here EFA/CFS produces factors based on correlation or variances explained in terms of percentage of factors. The mean score gives the level of each factor in terms of satisfaction, awareness or perception.
c) To perform bivariate or multivariate analysis on dependent variable established in factor analysis. This will determine the factors obtained as dependent variable if it differs on the basis of gender or ethnicity as independent variables requiring a bivariate or multivariate analysis. The statistical test involved can be independent t-test and ANOVA depending on the number of independent variable irrespective of the factors are obtained from the EFA or CFA.
d) To perform predictive analysis using a model for dependent factors. Here the factors obtained will be independent and dependent variable. However, this will require a questionnaire design structured with defined construct ideally demarcated into one dependent variable. Factors obtained can be EFA or CFA. Here the statistical methods used will be regression model.

15) Result and Exercise to Perform:

A flow chart developed for validation portrays the entire process of creation, validation and usability of questionnaire. A fictitious data created to write this paper has a number of table generated on analysis in Excel and SPSS as Content Validity Index (CVI), Reliability Indices (Cronbach’s alpha and Correlation Coefficient) and a number of Factor Analysis tables (see figure 1-4 and tables 1-8).

3. Analysing the Data for Interpretation:

The fictitious data has been created and Microsoft Excel data sheet developed as the raw data and is used for analysis and interpretation based on the objective of study, which is the validation of newly designed questionnaire based survey using 4-point Likert scale. Excel function used will determine the percentage weighting of each item on a scale from strongly agreed to strongly disagreed. However, based on objective and the hypothesis a questionnaire can be designed and developed with the intent to compare mean, perform linear regression and predictive statistics using t-test, ANOVA and regression methods. ANOVA can further be used beyond one-way ANOVA to multifactorial ANOVA and ANCOVA depending on research design developed to have dependent and independent variables and cofactors.

Here we consider analysis and interpretation of a questionnaire developed to measure the entrepreneurship abilities of participants after attending a training programme. This will use Microsoft Excel raw data worksheet either directly downloaded from the online survey or manually developed worksheet from the paper based survey. The analysis will use Excel functions (formula) for calculation and few steps to tabulate the data on another worksheet and to create chart in a graphic manner (see table 10 and figure 5). Following are the steps to analyse percentage value of Likert scale for its validity and reliability. The steps also include how to develop graph for its eye ball rolling evaluation of result. However, since the advent of Likert scale in 1932, there have been debates among the users about its best possible usability in term of reliability and validity of number of points on the scale[12, 13].

Step 1: Calculate the number of subjects or respondents using the following Excel function.

1) Select a cell and type Count: In the next cell in same row type =Counta (scroll the entire range in respective column as question) press enter. This will yield total number of responses.
2) Below the cell 1 above type, Count not responded: Type = Countblank (scroll the entire column range) and press enter.
3) Next to cell 2 below type Total count: Type = Sum (click 1 + click 2 outcome)

Step 2: For this step select cells and type Strongly agree, Agree, Disagree and Strongly disagree in a column below the total in above step and use following step for each criteria in Likert scale from Strongly agree to Strongly disagree.

1) In the cell below the Total type Strongly agree followed by Agree, Disagree and Strongly disagree one after another.
2) Start calculating from first criteria of Strongly agree: Type = Countif (scroll the entire column range using constant ($), "Strongly agree") and enter. This will give the numeric number of responses as strongly agree.
3) Drag the cell to fill in the rest of three criteria in cells below and later drag these 4 cells to fill in all the cells in four rows to get the numeric numbers for each question and each option (criteria.
4) Now to convert the entire data produced in decimals, click percentage (%) option in the menu bar on top to give percentage weighting of responses for all 4 options in the Likert scale.

Step 3: Now it is time to tabulate the data in a new worksheet to develop percentage weighting of each option in each question in their respective construct. Following are the steps to create new table.

1) Highlight the entire 4 rows in the worksheet and select a cell in new worksheet and click paste to select paste special followed by values and transformation and this will show a table with all the data. Selecting another cell before the data table and type question or item and below that enter Q1-Q5 (in current example) in each construct.
2) Now the table is ready for developing the graph by clicking insert on top menu bar and next choose bar chart showing option of 2-dimension and click to get the data in a graphic manner. Use edit option to type construct name on the top of the graph and readjust the font size as per the requirement.

4. Conclusion

Current paper described, how to create a questionnaire based survey-research and its validation in detail, which may help readers and new researchers to understand descriptive type of research using a survey design. It may help faculty interested to develop their research skills in questionnaire-based survey research specially during the COVID-19 pandemic to implement their innovative ideas with a concrete research outcome and sharing of experiences with global community. Multiple subject areas like perception of online teaching and its challenges, technology in teaching and learning, open-book exam, clinical education, assessment of competency, emergency remote learning and assessment and stresses of learning during COVID-19 pandemic and many more challenging situation encountered during COVID pandemic can be researched using newly created and validated questionnaires.

References

[1] Shahid Hassan. Teaching and Learning Approaches in Clinical Education during COVID-19 Pandemic: Time to Take-on Assessment as Learning. International Journal of Science and Research; Volume 9, Issue 11, November 2020: 1171-1174.
[2] Joshi A., Kale S. , Chandel S., Pal D. K. Likert Scale: Explored and Explained. British Journal of Applied Science & Technology 7(4): 396-403, 2015, Article no.BJAST.2015.157 ISSN: 2231-0843.
[3] Singh YK. Fundamental of Research Methodology and Statistics. New Delhi: Newage International (P) Ltd. Publisher; 2006.
[4] Culbertson JM. What is an attitude? Journal of Cooperative Extension. 1963; 79-84.
[5] Schwarz N, Bohner G. The construction of attitudes. In: Tesser A & Schwarz N, Editors. Blackwell Handbook of Social Psychology. Oxford: Blackwell. 2001;436- 457. 11.
[6] Polit D. F., Beck C. T. The content validity index: Are you sure you know what's being reported? Critique and recommendations. Research in Nursing and Health; 29 (5): 489-97.
[7] Yusoff M.S.B. ABC of content validation and content validity index calculation. Education in Medicine Journal. 2019;11(2):49–54. https://doi.org/10.21315/eimj2019.11.2.6
[8] Pituch, K. A. and Stevens, J., Applied multivariate statistics for the social sciences: Analyses with SAS and IBM's SPSS (6th ed.), Taylor & Francis, New York, 2016.
[9] Hair, J. J., Black, W.C., Babin, B. J., Anderson, R. R., Tatham, R. L., Multivariate data analysis, Upper Saddle River, New Jersey, 2006.
[10] Pallant, J., SPSS survival manual: a step by step guide to data analysis using SPSS, Open University Press/ Mc Graw-Hill, Maidenhead, 2010.
[11] Hair, J.F., Anderson, R.E., Tatham, R.L., and Black, W.C., Multivariate data analysis (5th ed.), N J: Prentice-Hall, Upper Saddle River, 1998.
[12] Colman AM, Norris CE, Preston CC. Comparing rating scales of different lengths: Equivalence of

scores from 5- point and 7-point scales. Psychological Reports. 1997;80:355-362.

[13] Preston CC, Colman AM. Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent Preferences. Acta Psychologica. 2000; 104;1-15.

Exercises Using Microsoft Excel and SPSS:

Please refer to steps, how to perform each method provided in the text along the appropriate site in method section and match it with appropriate table and figure for better understanding and follow through to do the take home exercises (task) as following.

Exercise 1: Create a fictitious data of 30 participants and 10 variables (questions/items) on a questionnaire based survey on a Likert scale of Strongly disagree = 1 to Strongly agree = 4

Exercise 2: Convert the text data to numeric numbers using recoding method and clean the data of unengaged participants using standard deviation statistics and subsequently transfer the file to SPSS worksheet

Exercise 3: On SPSS, identify the missing values and determine the normality of distribution and also perform the reliability test of Cronbach's alpha, Pearson and Spearman's coefficient.

Exercise 4: On SPSS worksheet perform the factor analysis choosing the right options of statistics, extraction and rotation and determine the factors extracted and number of variable assigned to those factors.

Exercise 5: On Microsoft Excel analyse and interpret the data using percentage weighting of Likert scale options for each variables (items) and develop the 2-dimension graphic representation.

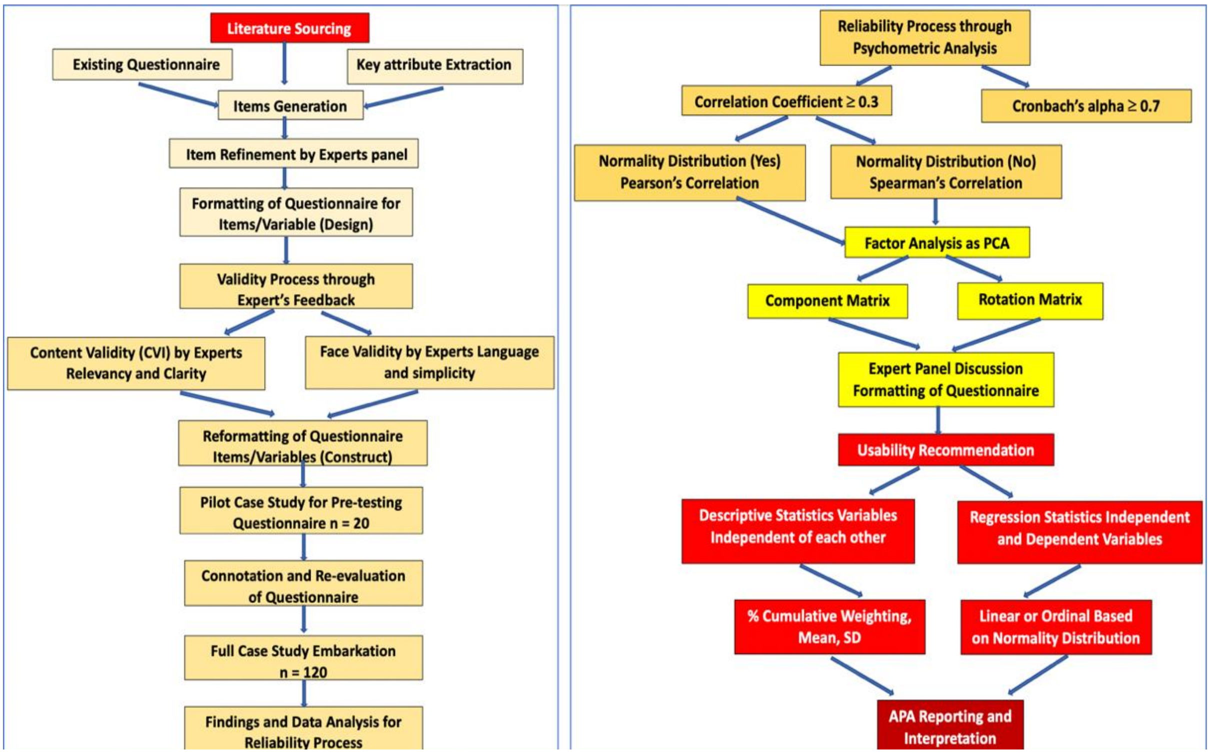

Figure 1: Flow chart of validation process from design to its usability in research involving reliability, validity and factor analysis

ISSN: 2319-7064
SJIF (2019): 7.583

1 = Not relevant. 2 = Somewhat relevant. 3 = Quite relevant. 4 = Highly relevant							
No	Item Description	Relevance					
		Ratter's Response					
		1	2	3	4	5	6
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							

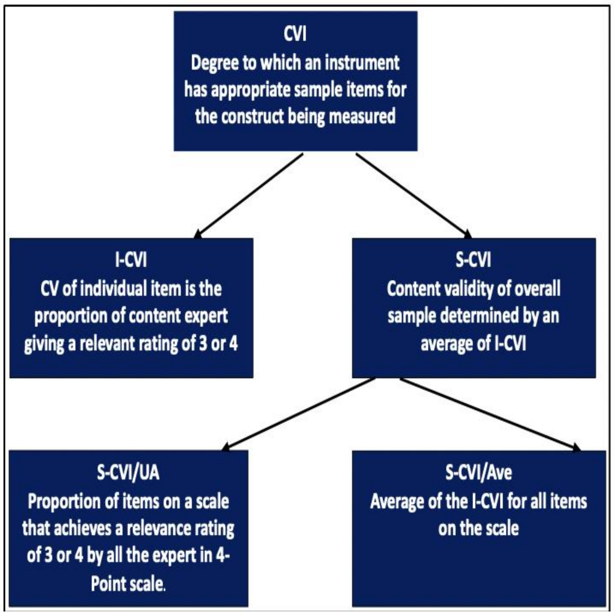

Figure 2: CVI classification (Polit and Beck, 2006) and ratting form provided for experts' opinion with description of construct

Table 1: Calculation of Content Validity Index as I-CVI, S-CVI/Ave and S-CVI/UA based on expert rating of items

Cluster/Item No	Rater 1	Rater 2	Rater 3	Rater 4	Rater 5	Rater 6	Agreement	I-CVI
Cluster1/Item1	4	4	1	3	4	3	5	0.83
Cluster1/Item2	3	4	3	3	4	4	6	1.00
Cluster1/Item3	3	4	3	4	4	3	6	1.00
Cluster1/Item4	3	4	3	4	4	3	6	1.00
Cluster1/Item5	3	4	1	4	4	4	5	0.83
Cluster1/Item6	2	4	2	4	4	4	4	0.67
Count	6					S-CVI/Ave		0.89
No of Rater	6					Total Agreement		3.00
						S-CVI/UA		0.50
	Rater 1	Rater 2	Rater 3	Rater 4	Rater 5	Rater 6	Agreement	I-CVI
Cluster2/Item1	4	4	4	3	4	4	6	1.00
Cluster2/Item2	4	4	3	3	4	3	6	1.00
Cluster2/Item3	3	4	3	3	4	4	6	1.00
Cluster2/Item4	3	4	1	3	4	4	5	0.83
Cluster2/Item5	3	4	4	3	4	4	6	1.00
Cluster2/Item6	3	4	1	3	4	3	5	0.83
						S-CVI/Ave		0.94
						Total Agreement		4.00
						S-CVI/UA		0.67
	Rater 1	Rater 2	Rater 3	Rater 4	Rater 5	Rater 6	Agreement	I-CVI
Cluster3/Item1	4	4	4	3	4	4	6	1.00
Cluster3/Item2	4	4	4	3	4	3	6	1.00
Cluster3/Item3	3	4	1	1	4	4	4	0.67
Cluster3/Item4	4	4	1	4	4	4	5	0.83
Cluster3/Item5	3	4	2	4	4	3	5	0.83
Cluster3/Item6	4	4	2	4	4	4	5	0.83
						S-CVI/Ave		0.86
						Total Agreement		2.00
						S-CVI/UA		0.33

Table 2: Reliability coefficient as Cronbach's alpha and Cronbach's alpha if item deleted of each item in cluster

Item Total Statistics (Construct: Prior Knowledge)				
5 items Chronbach's alpha =.783 All 15 Items Cronbach's alpha =.833	Prior Knowledge	Scale mean if item deleted	Corrected item-total correlation	Cronbach's alpha if item deleted
	Q1	11.80	.557	.745
	Q2	11.84	.476	.770
	Q3	11.93	.532	.761
	Q4	11.85	.632	.720
	Q5	11.85	.640	.719

Item Total Statistics (Construct: Motivation)				
5 items Chronbach's alpha =.757 All 15 Items Cronbach's alpha =.833	Prior Knowledge	Scale mean if item deleted	Corrected item-total correlation	Cronbach's alpha if item deleted
	Q6	11.94	.594	.687
	Q7	11.66	.517	.716
	Q8	11.91	.426	.750
	Q9	11.66	.610	.691
	Q10	11.65	.501	.722

Item Total Statistics (Construct: Competency)				
5 items Chronbach's alpha =.802 All 15 Items Cronbach's alpha =.833	Prior Knowledge	Scale mean if item deleted	Corrected item-total correlation	Cronbach's alpha if item deleted
	Q11	11.61	.570	.771
	Q12	11.61	.652	.742
	Q13	11.00	.450	.802
	Q14	11.33	.607	.758
	Q15	11.40	.656	.742

Table 3: Correlation Coefficient as Pearson's or Spearman's depending on normality of distribution in this case is Spearman's Correlation Coefficient since normality of data distribution was not achieved

Correlation Spearman's (rho)				
		Prior Knowledge	Motivation	Competency
Prior Knowledge	Correlation Coefficient Sig. (2-tailed) N	1 - 89	.655** <.001 89	.544** <.001 89
Motivation	Correlation Coefficient Sig. (2-tailed) N	.655** <.001 89	1 - 89	.622** - 89
Competency	Correlation Coefficient Sig. (2-tailed) N	.544** <.001 89	1 .622** 89	1 - 89
**Correlation significant at the .001 level (2-tailed)				

Paper ID: SR21512174451 DOI: 10.21275/SR21512174451

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Mean_Prior_Knowledge	.218	89	<.001	.921	89	<.001
Mean_Motivation	.164	89	<.001	.961	89	.009
Mean_Competency	.131	89	<.001	.963	89	.012

a. Lilliefors Significance Correction

Null Hypothesis for Normality Distribution

Construct 1, 2 and 3 are normally distributed

Alternative Hypothesis for Normality Distribution

Construct 1, 2 and 3 are not normally distributed

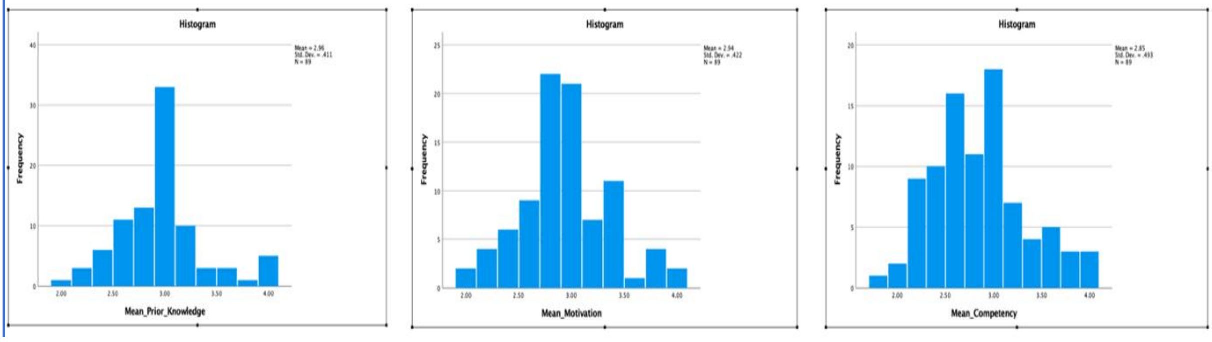

Figure 3: Test of normality using Shapiro-Wilk test assuming null hypothesis rejected here and the bar graph items mean score of three components (constructs)

Table 4: Component Correlation Matrix and KMO and Bartlett's test for qualifying the assumptions of factor analysis for intercorrelation component and data adequacy

Component Correlation Matrix

Component	1	2	3	4
1	1.000	.273	-.316	-.437
2	.273	1.000	-.235	-.329
3	-.316	-.235	1.000	.281
4	-.437	-.329	.281	1.000

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.876
Bartlett's Test of Sphericity	Approx. Chi-Square	536.329
	df	105
	Sig.	<.001

Table 5: Correlation matrices with Anti-image Covariance and Anti-image Correlation between the items to determine the measure of sampling adequacy in PCA

Anti-image Matrices

		Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15
Anti-image Covariance	Q1	.535	-.167	.004	-.113	-.027	-.004	-.144	-.062	.100	-.046	.050	.016	-.010	.033	-.102
	Q2	-.167	.562	-.061	.059	-.033	-.131	.022	.147	-.104	-.051	-.055	-.009	-.015	-.035	.055
	Q3	.004	-.061	.635	-.170	-.036	-.027	-.022	-.119	.011	-.024	-.014	-.005	-.008	.024	.001
	Q4	-.113	.059	-.170	.454	-.113	-.002	-.049	.118	-.086	.079	-.014	-.066	.061	-.072	-.024
	Q5	-.027	-.033	-.036	-.113	.446	-.127	-.101	.024	-.092	.008	-.023	.025	.031	.006	-.007
	Q6	-.004	-.131	-.027	-.002	-.127	.483	.026	-.101	-.036	-.122	.009	.041	-.029	-.061	-.001
	Q7	-.144	.022	-.022	-.049	-.101	.026	.509	-.101	-.028	-.071	.006	.062	-.044	-.011	-.082
	Q8	-.062	.147	-.119	.118	.024	-.101	-.101	.632	-.154	.082	-.084	-.092	.045	-.054	.055
	Q9	.100	-.104	.011	-.086	-.092	-.036	-.028	-.154	.480	-.051	.054	-.079	.014	-.037	-.032
	Q10	-.046	-.051	-.024	.079	.008	-.122	-.071	.082	-.051	.552	-.152	-.056	.003	-.037	.020
	Q11	.050	-.055	-.014	-.014	-.023	.009	.006	-.084	.054	-.152	.493	-.163	.061	-.034	-.097
	Q12	.016	-.009	-.005	-.066	.025	.041	.062	-.092	-.079	-.056	-.163	.459	-.065	-.009	-.125
	Q13	-.010	-.015	-.008	.061	.031	-.029	-.044	.045	.014	.003	.061	-.065	.660	-.236	-.091
	Q14	.033	-.035	.024	-.072	.006	-.061	-.011	-.054	-.037	-.037	-.034	-.009	-.236	.508	-.061
	Q15	-.102	.055	.001	-.024	-.007	-.001	-.082	.055	-.032	.020	-.097	-.125	-.091	-.061	.477
Anti-image Correlation	Q1	.845[a]	-.305	.006	-.229	-.055	-.009	-.276	-.106	.198	-.084	.098	.033	-.016	.063	-.202
	Q2	-.305	.842[a]	-.103	.117	-.067	-.252	.042	.246	-.200	-.092	-.104	-.017	-.025	-.065	.106
	Q3	.006	-.103	.915[a]	-.316	-.067	-.049	-.038	-.188	.019	-.040	-.025	-.009	-.012	.043	.002
	Q4	-.229	.117	-.316	.848[a]	-.250	-.004	-.102	.221	-.184	.158	-.029	-.144	.112	-.150	-.051
	Q5	-.055	-.067	-.067	-.250	.913[a]	-.273	-.212	.046	-.199	.017	-.049	.056	.057	.012	-.014
	Q6	-.009	-.252	-.049	-.004	-.273	.901[a]	.052	-.183	-.075	-.237	.018	.087	-.052	-.123	-.002
	Q7	-.276	.042	-.038	-.102	-.212	.052	.905[a]	-.178	-.057	-.135	.013	.127	-.075	-.021	-.166
	Q8	-.106	.246	-.188	.221	.046	-.183	-.178	.741[a]	-.279	.138	-.151	-.170	.069	-.095	.101
	Q9	.198	-.200	.019	-.184	-.199	-.075	-.057	-.279	.891[a]	-.100	.112	-.169	.025	-.074	-.066
	Q10	-.084	-.092	-.040	.158	.017	-.237	-.135	.138	-.100	.888[a]	-.292	-.112	.005	-.069	.039
	Q11	.098	-.104	-.025	-.029	-.049	.018	.013	-.151	.112	-.292	.866[a]	-.342	.107	-.069	-.201
	Q12	.033	-.017	-.009	-.144	.056	.087	.127	-.170	-.169	-.112	-.342	.873[a]	-.118	-.018	-.268
	Q13	-.016	-.025	-.012	.112	.057	-.052	-.075	.069	.025	.005	.107	-.118	.804[a]	-.408	-.161
	Q14	.063	-.065	.043	-.150	.012	-.123	-.021	-.095	-.074	-.069	-.069	-.018	-.408	.899[a]	-.124
	Q15	-.202	.106	.002	-.051	-.014	-.002	-.166	.101	-.066	.039	-.201	-.268	-.161	-.124	.906[a]

a. Measures of Sampling Adequacy(MSA)

Table 6: Descriptive statistics for normality of distribution checked skewness divided by standard error and looking for values within the range of -1.96 to +1.96 and the Communalities as initial and extracted values using Principle Component Analysis

Component	Descriptive Parameters	Statistics	Std. Error
Mean Prior Knowledge	Mean	2.9640	.04354
	95% CI for Mean Lower Bond	2.8775	
	Upper Bond	3.0506	
	Variance	.169	
	Std. Deviation	.41072	
	Skewness	.570	.255
	Kurtosis	.988	.506
Mean Motivation	Mean	2.9416	.04477
	95% CI for Mean Lower Bond	2.8526	
	Upper Bond	3.0305	
	Variance	.178	
	Std. Deviation	.42233	
	Skewness	262	.255
	Kurtosis	.266	.506
Mean Competency	Mean	2.8472	.05223
	95% CI for Mean Lower Bond	2.7434	
	Upper Bond	2.9510	
	Variance	.243	
	Std. Deviation	.49269	
	Skewness	.431	.255
	Kurtosis	-.186	.506

Communalities		
Items	Initial	Extraction
Q1	1.000	.607
Q2	1.000	.727
Q3	1.000	.518
Q4	1.000	.680
Q5	1.000	.672
Q6	1.000	.661
Q7	1.000	.605
Q8	1.000	.598
Q9	1.000	.577
Q10	1.000	.656
Q11	1.000	.640
Q12	1.000	.696
Q13	1.000	.729
Q14	1.000	.632
Q15	1.000	.667

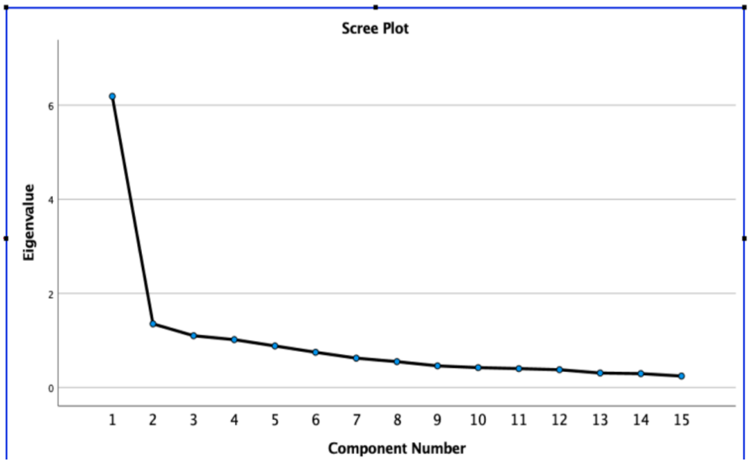

Figure 4: Scree Plot keeping with Eigen value of 1 indicates 4 components extracted of questionnaire designed with 3 constructs initially

Table 7: Total variance explained as Eigen value, initial extracted sum square and rotated sum square loading.

Total Variance Explained									
	Initial Eigen Values			Extraction Sums of Square Loading			Rotation Sums of Square Loading		
Comp	Total	%Var	Com%	Total	%Var	Com%	Total	%Var	Com%
1	6.189	41.261	41.261	6.189	41.261	41.261	3.074	20.494	20.494
2	1.355	9.033	50.294	1.355	9.033	50.294	2.312	15.415	35.909
3	1.102	7.345	57.639	1.102	7.345	57.639	2.253	15.018	50.927
4	1.019	6.794	64.433	1.019	6.794	64.433	2.026	13.506	64.433
5	.885	5.900	70.333						
6	.751	5.006	75.339						
7	.626	4.174	79.513						
8	.552	3.679	83.192						
9	.462	3.080	86.272						
10	.425	2.832	89.104						
11	.403	2.688	91.791						
12	.380	2.533	94.324						
13	.310	2.064	96.389						
14	.296	1.976	98.365						
15	.245	1.635	100.000						

Table 8: Component and rotation matrix with extraction method of principal component analysis using Varimax in Orthogonal rotation

Items	Extracted Component			
	1	2	3	4
Q5	.724	-.381		
Q9	.714			
Q15	.703			.321
Q6	.701			-.362
Q7	.686			
Q14	.680	.310		
Q4	.678	-.315		.323
Q12	.663	.459		
Q11	.644	.383		
Q10	.640			-.448
Q1	.612	-.368		
Q3	.601			
Q2	.593		.386	-.411
Q8	.480		-.578	
Q13	.445	.430	.500	.311

Items	Rotated Component			
	1	2	3	4
Q4	.775			
Q7	.697			
Q5	.684		.402	
Q1	.681			
Q3	.615	.339		
Q8		.737		
Q12		.685		.427
Q11		.655	.346	
Q9	.410	.513	.368	
Q2			.791	
Q10		.332	.696	
Q6	.352		.687	
Q13				.843
Q14				.647
Q15	.396	.363		.610

Table 9: Questionnaire initially created with 3 constructs readjusted after component extraction and rotation matrix suggesting 4 components (construct) with reshuffling of items

Item No	Construct I: Prior knowledge before attending the Entrepreneurship Training.
1	I have substantial knowledge and skills to understand entrepreneurship.
2*	Current training will boost my motivation and the existing knowledge in entrepreneurship.
3	Based on my entrepreneurship knowledge, I am confident to start my business in the area relevant to my profession.
4	I feel, my introduced products will be acceptable in the market and will compete fairly well with those in market.
5	I was capable of launching my project even prior to attending training program in entrepreneurship.
	Construct II: Motivation Acquired After Attending the Entrepreneurship Training.
6	I feel motivated to transform myself from an ordinary individual to a successful businessman.
7*	I have the inspiration of an entrepreneurs, known for their tenacity and commitment.
8	I have strong inspiration for entrepreneurship to continue even during the difficult part of process.
9	I know the importance of training has an impact on sustaining the growth and capabilities as entrepreneur.
10	I have a driving motivation that has changed the career of many self-made individuals in entrepreneurship.
	Construct III: Competency Developed to Practice Entrepreneurship in Near Future.
11	I have developed adequate knowledge and skills after entrepreneurship training to start my business
12	I admit to have wide gaps in knowledge and skills overcome after attending the entrepreneurship training.
	Construct IV: Confidence Developed to Practice Entrepreneurship in Near Future
13	I can now confidently consider to plan my own entrepreneurship project to practice.
14	I am sure after acquiring knowledge and skills my business will successfully compete with other business in market.
15	Current training has enabled me to think of setting up of a possible business to start soon.

Item No	Construct I: Prior knowledge before attending the Entrepreneurship Training.
1	I have substantial knowledge and skills to understand entrepreneurship.
2 (7)*	I have the understanding of an entrepreneurs, known for their tenacity and commitment.
3	Based on my entrepreneurship knowledge, I am confident to start my business in the area relevant to my profession.
4	I feel, my introduced products will be acceptable in the market and will compete fairly well with those in market.
5	I was capable of launching my project even prior to attending training program in entrepreneurship.
	Construct II: Motivation Acquired After Attending the Entrepreneurship Training.
6	I feel motivated to transform myself from an ordinary individual to a successful businessman.
7 (2)*	Current training will boost my motivation and the existing knowledge in entrepreneurship.
8 (10)	I have a driving motivation that has changed the career of many self-made individuals in entrepreneurship.
	Construct III: Competency Developed to Practice Entrepreneurship in Near Future.
9 (8)*	I have strong inspiration for entrepreneurship to continue even during the difficult part of process.
10 (9)*	I know the importance of training has an impact on sustaining the growth and capabilities as entrepreneur.
11	I have developed adequate knowledge and skills after entrepreneurship training to start my business.
12	I admit to have wide gaps in knowledge and skills overcome after attending the entrepreneurship training.
	Construct IV: Confidence Developed to Practice Entrepreneurship in Near Future
13	I can now confidently consider to plan my own entrepreneurship project to practice.
14	I am sure after acquiring knowledge and skills my business will successfully compete with other business in market.
15	Current training has enabled me to think of setting up of a possible business to start soon.

Table 10: Interpretation of result of Likert scale data analysis collected for validation of questionnaire Excel.

Question	Total Respondent	Strongly agree	Agree	Disagree	Strongly disagree	Total
Construct: Prior Knowledg						
Q1	89	13%	75%	11%	0%	100%
Q2	89	15%	69%	17%	0%	100%
Q3	89	15%	63%	22%	0%	100%
Q4	89	12%	72%	16%	0%	100%
Q5	89	11%	74%	15%	0%	100%
Construct:Motivation						
Q6	89	10%	57%	31%	1%	100%
Q7	89	19%	66%	15%	0%	100%
Q8	89	11%	57%	31%	0%	100%
Q9	89	15%	75%	10%	0%	100%
Q10	89	20%	66%	12%	1%	100%
Construct: Competency						
Q11	89	9%	49%	37%	4%	100%
Q12	89	10%	45%	43%	2%	100%
Q13	89	33%	58%	9%	0%	100%
Q14	89	16%	60%	25%	0%	100%
Q15	89	12%	60%	27%	1%	100%

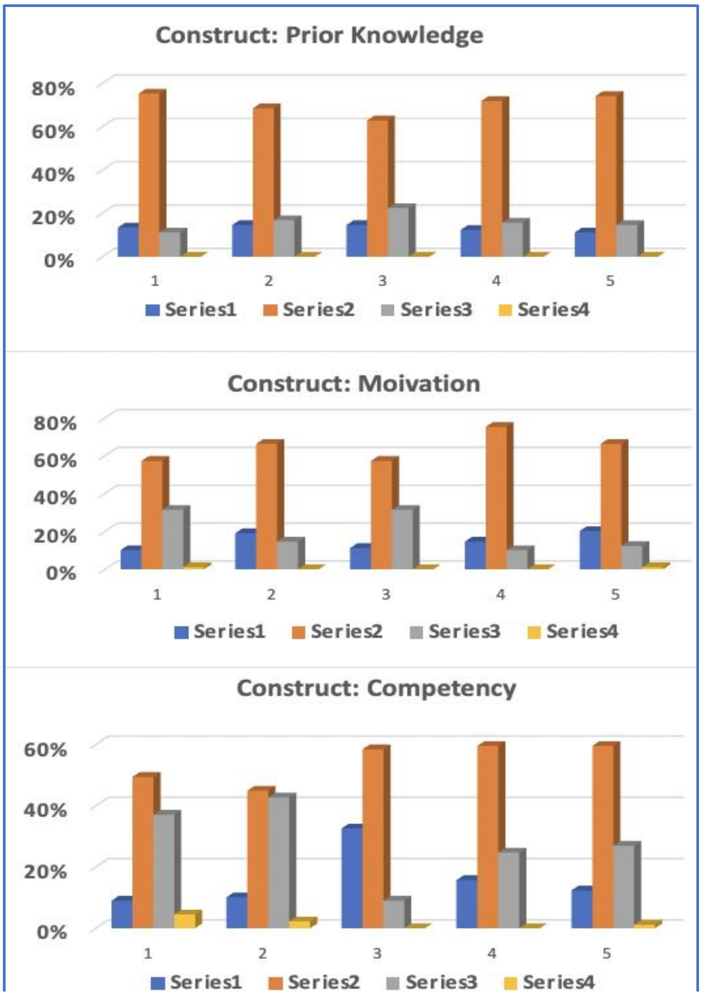

Figure 5: Result of Likert scale data analysis collected for validation of questionnaire using Microsoft Excel in a graphic manner

Appendix: Initially developed 15 items Questionnaire clustered into 3 constructs (components/factors) on a 5-point Likert scale from strongly agreed to strongly disagreed.

Entrepreneurship Training Impact (ETI) Inventory					
Item No	**Construct I: Prior knowledge before attending the Entrepreneurship Training.**	**Strongly Agree**	**Agree**	**Disagree**	**Strongly Disagree**
1	I have substantial knowledge and skills to understand entrepreneurship.				
2	Current training will be a simple refresher course for my existing knowledge in entrepreneurship.				
3	Based on my knowledge in entrepreneurship, I am confident to start my business in the area relevant to my profession.				
4	I feel, my introduced products will be acceptable in the market and will compete fairly well with those in market.				
5	I was capable of launching my project even prior to attending training programme in entrepreneurship.				
	Construct II: Motivation Acquired After Attending the Entrepreneurship Training.	**Strongly Agree**	**Agree**	**Disagree**	**Strongly Disagree**
6	I feel motivated to transform myself from an ordinary individual to a successful businessman.				
7	I have developed the motivation of an entrepreneurs, known for their tenacity and commitment.				
8	I have developed strong motivation for entrepreneurship to continue even during the difficult part of process.				
9	I know the importance of motivation has an impact on sustaining the growth and capabilities as future entrepreneur.				
10	I have a driving motivation that has changed the career of many self-made individuals in entrepreneurship.				
	Construct III: Competency Developed to Practice Entrepreneurship in Near Future.	**Strongly Agree**	**Agree**	**Disagree**	**Strongly Disagree**
11	I have developed adequate knowledge and skills after attending the entrepreneurship training to start my business/dental clinic.				
12	I admit to have wide gaps in knowledge and skills about entrepreneurship that I have overcome after attending the entrepreneurship training.				
13	I can now confidently consider to plan my own entrepreneurship project to practice.				
14	I am sure after acquiring knowledge and skills my business will successfully compete with other business in market.				
15	Current training has enabled me to think of setting up of a possible business to start soon.				