MATERIAL and METHODS

The Material for the study consisted of the entire batch of the II MBBS 2nd term students (n=120) who had passed the first MBBS examination in the first attempt and have completed one term of six months of clinical postings in the hospital.

INCLUSION CRITERIA

- Admitted to MBBS course in August 2008, of KLE University.
- Pass in 1st MBBS exam of August 2009, in the first attempt
- Completed a minimum of 6 months of Clinical Postings

Outline of the Study

Baseline: Computer Literacy, Emotional Intelligence
Tools of Assessment given to all, well prior; maintenance of confidentiality amongst group to prevent contamination

Primary Objective
Development of competence of Critical Thinking

Group-I
EBM “n=40”
Intervention

Group-II
Email “n=40”

Group-III
Traditional Control

Instructional Method

Outcome Measures / Tools
Competence of critical Thinking – CCTD
Change in Attitude – Alan’s EBM Survey
An informed consent was obtained from all the student participants. The aim of the study was explained to the student participants and the participation was entirely voluntary. They were emphasized that any of them were at liberty to leave the study at any stage of the study without giving any reason, if they find it inconvenient to continue to participate in the study.

The ethical clearance was obtained from the institutional review committee.

The total number of students enrolled in the study were only 118, as two students who remained absent on health grounds for initial introducing classes were excluded from the study.

All the student participants were divided into 3 groups’ viz., Group–EBM, Group-Email and Group-Traditional by stratified sampling method.

Methodology of Grouping: The average percentage of marks obtained by each student in the 1st MBBS University examination was recorded and was taken into consideration, for grouping.

The students were categorized or classified into 3 classes as detailed below depending upon their % marks obtained in 1st MBBS examination.
**Method of Randomization**

Divide the Cohort into subgroups and sample each group proportionately

<table>
<thead>
<tr>
<th>n=120 (H+ I + A)</th>
<th>120 Envelopes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>40=EBM</td>
</tr>
<tr>
<td></td>
<td>40=Email</td>
</tr>
<tr>
<td></td>
<td>40=Traditional</td>
</tr>
</tbody>
</table>

- **High Performers** - those students who have scored 65% or more.
- **Intermediate Performers** – those students who have scored between 60 and 64.9% of marks.
- **Average Performers**: those students who have scored between 50 and 59.9% of marks.

Thus, out of total 118 students, who were enrolled for the study;

- **7** were High Performers, (HP)
- **43** were Intermediate Performers (IP) and
- **68** were average performers. (AP)

Randomization was carried out using envelope method, to divide all the 3 classes (viz., HP, IP and AP) of performers into 3 groups of the study viz.,
(Group–EBM, Group-Email and Group-Traditional), each group containing equal number of all 3 kinds of performers, as follows.

To start with a total of 40 each of EBM, Email and Traditional paper chits were prepared and were put into identical types of envelopes. Thus 3 sets 40 blinded envelopes, each set, containing 40 paper chits of Group–EBM, Group-Email and Group-Traditional were prepared and kept ready.

These 120 blinded envelopes were regrouped into 3 new sets of envelope for each of the three classes thus containing 7, 43 and 68 envelopes for High, Intermediate and Average Performers, respectively, as follows.

The first set (Set 1) meant for High Performers contained 7 envelopes which contained 2, 2 and 3 paper chits of EBM, Email and Traditional chits respectively. Similarly the second set (Set II) meant for Intermediate Performers, contained 43 envelopes consisting of 15, 14 and 14 of EBM, Email and Traditional chits respectively. The third set (Set III) meant for Average Performers contained 68 envelopes which consisted of 23, 23 and 22 chits of EBM, Email and Traditional groups, chits respectively, as follows.

All the 3 classes of performers were made to pick one envelope each, from the respective set of envelopes meant for their class.

Finally as a result of this stratified sampling method of randomization, 3 groups viz., Group-1 (EBM) Group-2 (Email), and Group-3 (Traditional) contained equal number of all 3 kinds of performers.
Thus, Group-1 (EBM) (n=40) contained 2 High Performers, 15 Intermediate Performers and 23 Average Performers, Group-2 (Email) (n=39) contained 2 High Performers, 14 Intermediate Performers and 23 Average Performers, and the Group-3 (Traditional) (n=39) contained 3 High Performers, 14 Intermediate Performers and 22 Average Performers respectively (as shown in the table).

<table>
<thead>
<tr>
<th></th>
<th>EBM Chits</th>
<th>E-mail Chits</th>
<th>Traditional Chits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set-1 n=7 (For HP)</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Set-II n=43 (For IP)</td>
<td>15</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>Set-III n=68 (For AP)</td>
<td>23</td>
<td>23</td>
<td>22</td>
</tr>
</tbody>
</table>

The participants in the Group-EBM were scheduled to undergo a course on EBM and were also to receive emails; the participants of Group-Email were expected to receive periodic emails and the participants of Group-Traditional were not to receive any of the interventions. It was emphasized to all the participants of the study that the results or outcome of this study won’t have any effect whatsoever, irrespective of their groups, on their regular MBBS course or the respective University examination. They were well informed that the participants by enrolling themselves in this study, were helping the University to carry out an useful educational research.
Methodology

All the student participants were requested to keep their intervention details confidential to themselves to avoid contamination between the groups, so that the effects of intervention are not affected by this, contamination.

Introduction of the study to all the Student Participants:

The entire batch students belonging to II MBBS 2nd term who were fulfilling the Inclusion Criteria of the study were informed through a circular requesting them to assemble in the Pathology lecture hall at 3pm for the introductory class on 18-1-2010. The HoD’s of Pathology, Microbiology and Forensic Medicine were also invited for this introductory class. The batch was given a congratulatory note for their success in 1st MBBS examination, and also for completing one successful term of clinical postings.

The students were told to introduce themselves. Students were informed about the educational study that has been planned to be conducted in the University and their voluntary participation for the good cause was expected. The study design was informed and two things were made very clear to all of them. Firstly, they were told that each one of them have to sign a consent form and they were also told that they are at a liberty to leave the study at any stage and at any time if they find it inconvenient, without quoting any reason. Secondly, they were also told very clearly that the results or outcomes of this study are not going to bear any effect what so ever on their University MBBS examinations.

The copies of 1st MBBS marks card of all the students were collected and the aggregate marks they obtained in the 1st MBBS examination was also noted. The email addresses (email ID’s) of all the students were also noted. Majority of
the students were found to have their own laptops/computer systems. Their basic computer literacy was also noted using the respective questionnaire (see appendix).

The Pre-Course questionnaires of both tools of measurements of the study viz., California Critical Thinking Disposition Inventory (CCTDI) Questionnaire and Alan’s EBM Survey Questionnaire were given to all the participants and they were made to answer them there only in the class.

All the students were divided into three Groups viz., EBM, Email and Traditional by the methodology described as above and all the students were made to note and remember the group they belong to. Group wise list of the students was also prepared and all the 3 groups viz., Group-I EBM, Group-II Email, and Group-III Traditional contained almost equal number of participants belonging to equal number of category/class as per their aggregate 1st MBBS scores. An email was sent to all the students informing of their Group, on the same evening and all were told to acknowledge the same.

**Intervention Plan of the Study**
Effect of Intervention

Effects of Modes of Instruction

**Sample Size of the Study**

When we design an educational study because larger samples result in small sampling errors, which means that our sample size will be closer to true population size. In the extreme case sampling error would be zero if we include complete population in our study rather than drawing a sample. As a rule of thumb it is recommended that the whole population be used when population number is 100 or less. Since the total number of students in the class / batch chosen for the study is 120, the entire batch has been enrolled for the study.

Sample size also depends upon random sampling method; more efficient the method is, smaller is the sample size required. The stratified random sampling method an efficient sampling method used in this study requires smaller sample size.

Another sampling strategy as described by Margaret Lecompte, Judith and Renatatesch (1993) is, known as ‘Criterion-based Selection’
researcher develops inclusion criteria to be used in selecting people, as is done in this study. The inclusion criteria for our study are explicitly defined.

Another well known qualitative researcher Michael Patton (1990) used the term ‘Purposeful Sampling’ to describe the same process because individuals or cases are selected that provide the information needed to address the purpose of the research. The terms ‘Criterion-based Selection’ and purposeful sampling are synonyms and both describe what is called as Purposive Sampling which is used both in qualitative as well as quantitative research.

The key point is that the researcher should pick a sample that can be used to meet the purpose of research study and answer research questions while meeting cost and other constraints. In other words a sample should be as large as the researcher can obtain with reasonable expenditure of time and energy.

There are few guidelines available also to suggest minimum number of subjects needed for different types of studies. For descriptive studies a minimum of 100, for correlational studies a minimum of 50 and for experimental study and casual comparative study a minimum of 30 individuals per group are essential. Sometimes experimental studies with only 15 individuals in each group can be defended if they are very tightly controlled; studies using only 15 subjects per group should probably be replicated.

Based on earlier analytical study done by Candan O et al, in undergraduate nursing students in the year 2007 using the CCTDI to compare the effects of PBL and Traditional education method (lecturing) on the student’s critical thinking levels, wherein the difference between the intervention group
and the control group was 10; we for our study predicted the difference between the scores in control group and intervention group to be about 15; we expected the difference to be much higher (about 1.5 times) than the previous study because we assumed that medical students will have or have to develop higher CT abilities than nursing students because of their higher intelligence status, or need.

Assuming the power of the study to 80%, the expected difference in scores between the groups to be about 15, and assuming the level of significance (i.e., \( \alpha \) value) to be 0.05, the calculated sample size using the formula

\[
N = \left( \frac{Z_{\alpha} + Z_{\beta}}{\delta} \right)^2 \frac{d}{n}
\]

where
- \( n \) = Sample size
- \( d \) = Mean difference
- \( \delta \) = Standard Deviation
- \( \alpha = 0.05 \)
- \( Z_{\alpha} = 1.96 \) for two sided
- 1.65 for one sided

Will be 40

Our chosen sample size in the study of 40 in each of the groups can also be justified on this basis.
**Methodology**

**Stratified Random Sampling:**

It is a process in which certain sub groups or strata are selected for the sample in the same proportion as they exist in the population. Here in our study the 3 sub groups or strata viz., High Performers, Intermediate Performers and Average Performers are distributed randomly and equally between all the 3 groups so that they exist in the same proportion in all the 3 groups as they exist in the entire class of students.

The advantage of stratified random sampling is that it increases the likelihood of representativeness, especially if one’s sample is not very large. It virtually ensures that any key characteristics of individuals in the population are included in the same proportions in the sample. The disadvantage is that it requires still more effort on the part of the researcher.

**Statistical Analysis of the Data**

There are two basic types of inference techniques that researchers use.

I) Parametric techniques make various kinds of assumptions about the nature of population from which samples involved in the study are drawn.

II) Non-Parametric techniques on the other hand make few (if any) assumptions about the nature of population from which samples are taken.

An advantage of parametric techniques is that they are generally more powerful than non-parametric techniques and hence are much more likely to reveal a true difference or relationship if one really exists. Their disadvantage is
that the researcher cannot satisfy the assumptions they require (for eg. that the population is normally distributed on the characteristic of interest). The advantage of non-parametric techniques is that they are simpler and safer to use when the researcher cannot satisfy the assumptions underlying the use of parametric tests.

The use of non-parametric methods may be necessary when data have a ranking but no clear numerical interpretation such as, when assessing preferences; in terms of levels of measurement for data on ordinal scale.

Due to their reliance on fewer assumptions non-parametric methods are more robust. The wider applicability and increased robustness on non-parametric tests comes with a cost in cases where parametric tests would be appropriate, non-parametric tests have less power. In other words larger samples size can be required to draw conclusions with the same degree of confidence.

However the thumb rule can be stated as follows:

Use the tests of statistical significance only to evaluate generalizability not to evaluate the magnitude of relationships.

With the explosion of information and communication technology in the field of biomedical statistical, calculation of statistics has become rather easy and quick and hence it is suggested to use both parametric and non-parametric tests to analyse the data. When the results are consistent, interpretation will thereby be strengthened. When the results are not consistent discuss possible reasons.
Methodology

Statistical Analysis: Statistical Tests: \(^{57}\)

Parametric Tests:

1) The ‘t’ Test: It is a parametric statistical test used to see whether the difference between the means of two samples is significant. The test produces a value for ‘t’, which the researcher then checks in a statistical table to determine the level of significance that has reached. If the 0.05 level of significance is reached, the null hypothesis is customarily rejected and concludes that a real difference exists.

Non-Parametric Tests:

1) **Kruskal-Wallis one way analysis of variance:**

   It is a non-parametric test for testing equality of population ‘Medians’ among groups. It is identical to a one way analysis of variance with data replaced by their ranks. It is an extension of the Mann-Whitney U test to 3 or more groups. Since it is a non-parametric method, the Kruskal-Wallis test does not assume a normal population unlike the analogous one way ANOVA. However the test assumes an identically shaped and scaled distribution for each group except for any difference in medians.

2) **Wilcoxon Signed-rank test:**

   It is a non-parametric statistical hypothesis test used when comparing two related samples or repeated measurements on a single
sample to assess whether their population means differ i.e., it is a paired difference test. It can be used as an alternative to the paired student’s ‘t’ test when population cannot be assumed to be normally distributed or the data is on the ordinal scale.

3) **Cronbach’s Alpha**:  

   It is generally used as a measure of internal consistency or reliability of a psychometric instrument. In other words it measures how well a set of variables or items measure a single, one dimensional latent aspect of individuals. Generally many quantities of interest in medicine such as anxiety or degree of handicap are impossible to measure explicitly. In such case we ask a series questions and combine the answers into a single numerical value.

   Cronch Bach’s Alpha generally increases when the correlations between the items increase. For this reason it is also called the internal consistency or the internal consistency reliability of the test. Cronbach’s alpha has been calculated to measure the reliability of AEBM ASQ.

   A value of Cronbach’s alpha of more than 0.6 is considered as acceptable and a value of more than 0.8 as excellent.

   Hence our study the data has been analysed and the level of significance has been calculated using both the Parametric as well as Non-Parametric tests.
Details of Intervention

Course on EBM for Group –I (EBM)

Duration – Activities spread over 8 months

i) Pre /Intra course reading materials through emails

ii) Two half day Seminars/Hands on Training workshops on EBM

iii) One interactive lecture of one hour duration

iv) One group activity of 1 hour duration

Details of Intervention

A total of five interventions were planned for Group-I(EBM) in a span of 8 months starting from 26th Feb. 2010 upto the end of October 2010. The time table for the interventions was prepared taking into consideration their regular time table of II MBBS program so that there will not be any conflict between their regular course of MBBS study and the study intervention (see table).
**Curricular Outline /Details**

<table>
<thead>
<tr>
<th>Sl No.</th>
<th>Activities</th>
<th>Date and Time</th>
</tr>
</thead>
</table>
| 1.     | Welcome and Introduction to the study for the entire class obtaining the consent for the study  
|        | Basic information/data- 1 MBBS marks, Computer Literacy, Emotional Intelligence, Pre course CCTDI Questionnaire, and Alan’s EBM Survey Questionnaire.  
|        | Division into 3 groups using envelope method and stratified sampling  
|        | Welcome E-mail to all 3 groups confirming their group after the session                                                                                                                                  | 18 Jan 2010 3pm to 5pm |
| 2.     | 1 Intervention: EBM Group- First Seminar on EBM…………………………………………………………...  
|        | Handouts of the seminar were distributed to all the participants of the Group-I (EBM).                                                                                                                      | 26 Feb 2010 3pm to 5pm |
| 3.     | 2 Intervention: An article on EBM was sent to EBM Group                                                                                                                                                | 15 April 2010         |
| 4.     | 3rd Intervention: Small group activity in Department of Medical Education (DOME)                                                                                                                                 | 25 May 2010 3 pm to 4 pm |
| 5.     | 4th Intervention: An interactive lecture on EBM An article on EBM was sent to all the participants                                                                                                                                                        | 28 Aug. 2010          |
| 6.     | 5th Intervention: Second seminar on EBM…………………………………………………………...  
|        | Handouts of the second seminar were distributed to all the participants.                                                                                                                                  | 22 Oct. 2010 2pm to 4 pm |
| 7.     | Summative(final) Assessment: Post Course CCTDI Questionnaire, Alan’s EBM survey for all the 3 groups                                                                                                    | 22 Oct. 2010 5pm to 6pm |
Methodology

All the interventions were carried out in the Dept. of Medical Education Hall, wherein different modes of instruction viz., Interactive lecture, Small group activity, Brain storming session, Hands on training, Data Literature search etc. could be carried out quite conveniently.

Details of a course on EBM-for Group-I (EBM)

The course on EBM consisted of two half day seminar cum hands on training workshops on Principles and Practice of EBM, (2 hours duration each) one interactive lecture (one hour duration) and one small group activity (one hour duration).

Pre seminar/activity reading material was emailed to all the participants of EBM group at least one week prior to the activity and all the students were requested to go through the same and come well oriented for the study activity.

Details of the activity, logistic details, and typical lesson plan are in the table below.
## Typical Lesson Plan for the Seminar cum Hands on Training on EBM

<table>
<thead>
<tr>
<th>Topic</th>
<th>Activity</th>
<th>Resources</th>
<th>Who</th>
<th>Duration</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction and Objectives</td>
<td>Pre-Seminar Questionnaire</td>
<td>Instructor, LCD, Laptop</td>
<td>PFK/JN</td>
<td>30 mins</td>
<td>1500-1530</td>
</tr>
<tr>
<td></td>
<td>Brainstorming</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Concept and scope</td>
<td>Interactive Lecture</td>
<td>Instructor, LCD, Laptop</td>
<td>PFK/JN</td>
<td>30 mins</td>
<td>1530-1600</td>
</tr>
<tr>
<td></td>
<td>Group activity and presentation</td>
<td>Instructor, LCD, Laptop</td>
<td>PFK/JN</td>
<td>40 mins</td>
<td>1600-1640</td>
</tr>
<tr>
<td></td>
<td>Summary feedback</td>
<td>Instructor, LCD, Laptop</td>
<td>PFK/JN</td>
<td>15 mins</td>
<td>1640-1655</td>
</tr>
<tr>
<td></td>
<td>Post seminar Questionnaire</td>
<td>Questionnaire</td>
<td>-</td>
<td>05 mins</td>
<td>1655-1700</td>
</tr>
</tbody>
</table>

PFK – Dr. P.F. Kotur  
JN – Dr. Jyoti Nagmoti

### Details of the Intervention for Group-II (Email).

All the participants belonging to Group-II (Email) were scheduled to receive the Pre-Course material, handouts etc through email on the same day on which an intervention was made for EBM group. Students were requested again to acknowledge the receipt of the email. However the information as to whether the email was opened or not by the respective student was also obtained through the options available with email providers.

The participants of Group-II(Email) were thus scheduled to receive all the Pre-course materials and handouts of all the 5 interventions of the EBM group.
Details of Intervention for Group-III(Traditional). There was no intervention of any kind (neither any activity/mode of instruction or email) to the participants of Group-III (Traditional).

Methods of Elimination of Errors in the Study

All the possible measures available to eliminate any kind of errors, those can creep into the study, at any stage, were incorporated in the study design as listed below: Thus, the study participants were divided into three groups viz Group-I(EBM), Group-II (Email) and Group-III (Traditional) each scheduled to receive a different kind of intervention. Comparison of the effects of the intervention between Group-I and Group-III and between Group-II and Group-III was supposed to yield information about the effect of respective intervention. Comparison between Group-I and Group-II was supposed to yield information about the effects of different modes of instruction.

1. Control group / Study group
2. Blinding-Code numbers to groups and participants
3. Stratified sampling
4. Maintenance of confidentiality
5. CCTDI-Copyright, No Key, Analysis Package, ICR answer sheet.

Tools of Assessment

The two measurable objectives of the study are the ability of the medical student to think critically whenever he / she encounters a medical dilemma and
the capability to justify the practice of EBM as an aid to clinical decision making process. The two tools selected for the study to measure these objectives are:

i) California Critical Thinking Disposition Inventory (CCTDI)

ii) Alan’s EBM Attitude Survey Questionnaire (AEBM ASQ)

Both of the tools have been in use in the field of qualitative research in education and the psychometric qualities of both of them have been tested and well documented.

i) California Critical Thinking Disposition Inventory (CCTDI)
Methodology

It is a premier, time tested tool for surveying the dispositional aspects of Critical Thinking since the year 1986. Ever since then it is being used extensively all over the world to assess the Students, Graduates and Professionals in various educational institutions in more than 25 countries in the World.

CCTDI is the premier critical thinking disposition instrument in the world today. Based on the Delphi expert consensus definition of Critical Thinking, description of the ideal critical thinker the instrument has been used in decades of empirical and conceptual studies of human reasoning behaviour.

Drs. Peter and Noreen Facione developed the CCTDI in the 1990’s. Since that time the instrument has been collaboratively translated to numerous culturally and psychometrically valid and reliable language forms that are now used by employer and educational institutions around the world. Independently the scores on CCTDI inform the test administrator of the learning readiness or hiring potential of the test taker. Results from the tests can be used to demonstrate the accomplishment of stated goals related to the development and demonstration of strength in critical thinking. The CCTDI is specifically designed to measure the disposition to engage problems and make decisions using critical thinking. The CCTDI measures the attitudes and values that measure that influence the test taker’s capacity to learn and to effectively apply critical thinking skills.
The CCTDI requires no educational preparation and its items can be easily read by the late adolescent or adult population. The instrument is not cognitively fatiguing and as a result can be administered at any time.

The instrument contains 75 Likert scale items chosen for their internal consistency and their ability to discriminate between the respondents. The seven scales of the CCTDI are composed of nine to 12 items with the scales interspersed throughout the instrument. Test takers are invited to indicate the degree to which they agree or disagree with each of the 75 items. The tool has
Cronbach’s alpha for total scale of 0.92. It has 7 subscales with Cronbach’s alphas for each of the subscales as follows:

- Truth seeking 12 items, alpha 0.71
- Open-mindedness 12 items, alpha 0.73
- Analyticity 11 items alpha 0.72
- Systematicity 10 items alpha 0.74
- CT and Self confidence 10 items alpha 0.78
- Inquisitiveness 10 items alpha 0.80
- Maturity 10 items alpha 0.75

Total points from the seven scales determine an individual’s CT disposition.

Responses are recorded using a 6 point Likert scale ranging from “strongly agree” to “strongly disagree”. Respondents must agree or disagree with the idea being offered by each item (there is no neutral option) with each item prompt. Each idea (item) is consistent with or in opposition to a recognised critical thinking dispositional attribute or attitude. In an important respect CCTDI refines and extends the conceptualization of critical thinking in the Delphi Report. The CCTDI prompts express familiar opinions beliefs, values, expectations and perceptions. Phrased in standard english, the items use no technical vocabulary or critical thinking jargon. No specialized educational level or content knowledge is presumed.

It is protected by international copyrights of Insight Assessment Capscore™ California USA.
The CCTDI can be administered in about 20 minutes. The tool invites the respondents to indicate the extent to which they agree or disagree with 75 statements expressing beliefs, values, attitudes and intentions that relate to reflective formation of reasoned judgment after exercising the Critical Thinking ability on the issue concerned. The basic structure of the instrument is based on methodology of Likert scale.

Each of the 7 components of Critical Thinking viz.,

1) Truth Seeking Ability
2) Open Mindedness
3) Analytical Skills
4) Systematicity
5) Confidentiality
6) Inquisitiveness and
7) Maturity

are addressed and measured through the responses of the participants to the afore-mentioned 75 statements.

A person may be disposed towards Truth seeking or bias, towards Open-mindedness or Intolerance, towards being Analytical anticipating possible consequences or being heedless of them, towards proceeding in a systematic or unsystematic way, towards being confident in the powers of reasoning or mistrustful of thinking, towards being Inquisitive or resistant to learning, and towards Mature and nuanced judgment or toward rigid simplistic thinking. An overall critical thinking disposition score is also calculated.
The CCTDI measures above character logical attributes and its scale scores profile the survey respondent on these seven dimensions.

Higher scores on the CCTDI are positively correlated with a strong desire to apply one's critical thinking skills in decision making and problem solving and with the capacity to benefit from educational training. High scores on this overall measure indicate that the valuation of critical thinking as an approach to analyzing and resolving complex, novel, high stakes problems is a dominant habit of mind. Low scores on this overall measure indicate that the test taker probably has none of the attitudes and attributes associated with the strong Critical Thinker.

Over all scores below 240 should be considered as very low scores.

Individuals who score in this range don’t often seek additional educational facility or opportunity. Typically they exhibit the collection of attributes, under each of scale measures as being the opposite of the desired attribute.

Scores below 270 over all are rare in the undergraduate population.

Overall scores above 350 are relatively rare in the undergraduate population and when all scores are above 40 the test taker has demonstrated significant and generalised strength in critical thinking disposition.
**Methodology**

**CCTDI scale scores:**

In addition to providing CCTDI total score for each test taker, the CCTDI measures and reports seven attributes of CT disposition; these 7 scores are far more informative than the overall score. Some individuals will be seen to have strengths in some areas but weaknesses in other areas, and over all profile has implications for how amenable an individual is for training high order thinking skills. (HOTS) scales on each of the CCTDI scales range between 10 and 60. While one can analyze these scores as continuous data, noting point improvements over time the scores are also indicative of qualitative ranges as described below. Movement from one range to another is an indication of overall attitudinal change.

**Scale scores in the 10-29 range (Low):**

These scores are indicative of a lack of valuation, aversion to or even hostility toward the attitude or attribute being measured. Lower scores in this range indicate that the attitude or attribute is a negative habit of mind likely to affect thinking in most situations. While movement from this level of score to one demonstrating strength on the attribute has been seen in test takers after participation in an educational activity or opportunity aimed at the scale construct, it is more common to see individuals express more ambivalent valuation of the measure at the 2\textsuperscript{nd} testing date.
Scale Scores in the 30-40 range (Ambivalent):

These scores are indicative of ambivalent or inconsistent endorsement of the attitude or attribute being measured. Test takers frequently seen to move from this score to category to higher range as a result of completing an educational training program aimed at the educational training program aimed at the scale construct. More rarely such results also result in some individuals regressing to a more adverse or hostile range at 2nd testing. This movement towards aversion is consistent with other observations of attitude formation. Having the habit of mind of engaging life and work problems with ones HOTs in order to determine what to believe and what to do (Critical thinking) requires courage, persistence, honesty and organisation. It is reasonable that some individuals may reject the challenge or find this path too fearful or difficult. Many of these fall into non reflective strategies to address life and work problems scales scores in the range of 40-50 range (positive). These scores indicate consistent endorsement and valuation of attitude or attribute being measured.

Scales scores in the range 50-60 (High):

Scores in this range indicate that the attribute or attitude is a positive habit of mind and likely to factor into the individual’s approach to all high order thinking (reflective problem definition and problem solving) particularly when situation is of high consequence.

Research on CT by age suggests that midlevel scores are more typical of younger individuals who may still be evolving in terms of their cognitive development. Studies have shown considerable growth in individuals in this age
range coincidental to undergraduate programs aimed at improving HOTS and lifelong learning behaviour. Some theorize that exposure to mentors who model these dispositions has great potential to strengthen attitudes in this area.

ii) Alan’s EBM Attitude Survey Questionnaire (AEBM ASQ)

It is an instrument which measures the attitude of the person towards the principles and practice of EBM. This scale is again a time tested tool, with accepted Psychometric qualities and has been in use since 2007.

There are 25 questions to be answered by the student participants by conveying whether the test taker, strongly agrees/disagrees with the statement. Again it is based on Likert Scaling Methodology.

Evaluation of the Participants

- Pre course assessment of all 3 Groups
- Post course assessment of all 3 Groups

Both the questionnaires (tests) were administered to all the participants of all the three groups, both before any intervention was made to any of the group as well as after the interventions were administered in Group-I(EBM) and Group-II(Email).

Likert scale

A Likert scale, or more accurately a Likert-type scale, is a psychometric scale commonly used in questionnaires, and is the most widely used scale in survey research, such that the term is often used interchangeably
with rating scale even though the two are not synonymous.\(^{(59)}\) There is debate as to what is a true Likert scale and what is a 'Likert-type' scale. Likert's original scale (in his PhD thesis) was bipolar, with five points running from one extreme to another, through a neutral central position, ranging from 'Strongly Agree' to 'Strongly Disagree'.

When responding to a Likert questionnaire item, respondents specify their level of agreement to a statement. The Likert Scale is an ordered, one-dimensional scale from which respondents choose one option that best aligns with their view. It is a method of ascribing quantitative value to qualitative data, to make it amenable to statistical analysis. A numerical value is assigned to each potential choice and a mean figure for all the responses is computed at the end of the evaluation or survey. The scale is named after its inventor, the US organizational-behavior psychologist Dr. Rensis Likert (1903-81).

An important distinction must be made between a Likert scale and a Likert item. The Likert scale is the sum of responses on several Likert items. Because Likert items are often accompanied by a visual analog scale (e.g., a horizontal line, on which a subject indicates his or her response by circling or checking tick-marks), the items are sometimes called scales themselves. This is the source of much confusion; it is better, therefore, to reserve the term Likert scale to apply to the summated scale, and Likert item to refer to an individual item.

A Likert item is simply a statement which the respondent is asked to evaluate according to any kind of subjective or objective criteria; generally the
level of agreement or disagreement is measured. Often five ordered response levels are used, although many psychometrists advocate using seven or nine levels; a recent empirical study found that a 5- or 7-point scale may produce slightly higher mean scores relative to the highest possible attainable score, compared to those produced from a 10-point scale, and this difference was statistically significant. In terms of the other data characteristics, there was very little difference among the scale formats in terms of variation about the mean, skewness or kurtosis.

The format of a typical five-level Likert item is:

1. Strongly disagree
2. Disagree
3. Neither agree nor disagree
4. Agree
5. Strongly agree

**5-point traditional Likert scale:**

<table>
<thead>
<tr>
<th>I like going to Chinese restaurants</th>
<th>Strongly agree</th>
<th>Tend to agree</th>
<th>Neither agree nor disagree</th>
<th>Tend to disagree</th>
<th>Strongly disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
</tbody>
</table>
5-point Likert-type scale, not all labeled:

<table>
<thead>
<tr>
<th>Good</th>
<th>Neutral</th>
<th>Bad</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

When I think about Chinese restaurants I feel

6-point Likert-type scale:

<table>
<thead>
<tr>
<th>Never</th>
<th>Infrequently</th>
<th>Infrequently</th>
<th>Sometimes</th>
<th>Frequently</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

I feel happy when entering a Chinese Restaurant

Likert scaling is a bipolar scaling method, measuring either positive or negative response to a statement. Sometimes a four-point scale is used; this is a forced choice method since the middle option of "Neither agree nor disagree" is not available.

Likert scales may be subject to distortion from several causes. Respondents may avoid using extreme response categories (central tendency bias); agree with statements as presented (acquiescence bias); or try to portray themselves or their organization in a more favorable light (social desirability bias). Designing a scale with balanced keying (an equal number of positive and negative statements) can obviate the problem of acquiescence bias, since acquiescence on positively keyed items will balance acquiescence on negatively keyed items, but central tendency and social desirability are somewhat more problematic.
After the questionnaire is completed, each item may be analyzed separately or in some cases item responses may be summed to create a score for a group of items. Hence, Likert scales are often called summative scales.

Whether individual Likert items can be considered as interval-level data, or whether they should be considered merely ordered-categorical data is the subject of disagreement. Many regard such items only as ordinal data, because, especially when using only five levels, one cannot assume that respondents perceive all pairs of adjacent levels as equidistant. On the other hand, often (as in the example above) the wording of response levels clearly implies a symmetry of response levels about a middle category; at the very least, such an item would fall between ordinal- and interval-level measurement; to treat it as merely ordinal would lose information. Further, if the item is accompanied by a visual analog scale, where equal spacing of response levels is clearly indicated, the argument for treating it as interval-level data is even stronger.

When treated as ordinal data, Likert responses can be collated into bar/graph charts, central tendency summarised by the median or the mode (but some would say not the mean), dispersion summarised by the range across quartiles (but some would say not the standard deviation), or analyzed using non-parametric tests, e.g. chi square test, Mann Whitney test, Wilcoxon signed - rank test, or Kruskal-Wallis test. Parametric analysis of ordinary averages of Likert scale data is also justifiable by the Central Limit Theorem, although some would disagree that ordinary averages should be used for Likert scale data.
Responses to several Likert questions may be summed, providing that all questions use the same Likert scale and that the scale is a defendable approximation to an interval scale, in which case they may be treated as interval data measuring a latent variable. If the summed responses fulfill these assumptions, parametric statistical tests such as the analysis of variance can be applied. These can be applied only when more than 5 Likert questions are summed.

Data from Likert scales are sometimes reduced to the nominal level by combining all agree and disagree responses into two categories of "accept" and "reject". The chi square test, Cochran Q, or Mc Nemar test are common statistical procedures used after this transformation.

Consensus based assessment can be used to create an objective standard for Likert scales in domains where no generally accepted standard or objective standard exists. Consensus based assessment (CBA) can be used to refine or even validate generally accepted standards.

A benefit is that questions used are usually easy to understand and so lead to consistent answers. A disadvantage is that only a few options are offered, with which respondents may not fully agree.

As with any other measurement, the options should be a carefully selected set of questions or statements that act together to give a useful and coherent picture.
A problem can occur where people may become influenced by the way they have answered previous questions. For example if they have agreed several times in a row, they may continue to agree. They may also deliberately break the pattern, disagreeing with a statement with which they might otherwise have agreed. This patterning can be broken up by asking reversal questions, where the sense of the question is reversed. There is much debate about how many choices should be offered. An odd number of choices allow people to sit on the fence. An even number forces people to make a choice, whether this reflects their true position or not.

Some people do not like taking extreme choices as this may make them appear as if they are totally sure when they realize that there are always valid opposing views to many questions. They may also prefer to be thought of as moderate rather than extremist. They thus are much less likely to choose the extreme options.