

*Full Length Research Paper*

# Co-educational tutorial classes and their significance on gendered test scores of Wollo University students: A before-after analyses

Mu'uz Gidey

Wollo University, Ethiopia.

Received 17 May, 2015; Accepted 31 July, 2015

**This action research is carried out in a practical class room setting to devise an innovative way of administering tutorial classes to improve students' learning competence with particular reference to gendered test scores. A before-after test score analyses of mean and standard deviations along with t-statistical tests of hypotheses of second year Civics and Ethical Studies Students administering 18 hours tutorial classes had been carried out. The findings showed the hitherto upheld gender gap based view that the average female student scores lag behind their male counter parts. Test scores have improved significantly for female students than male after tutorial. The significance test as to whether tutorial classes have gendered correlates showed high pedigree, contrary to common expectations, towards female students indicating the need for rearticulating gender sensitivity to the needs of male students. The study has shown tutorial classes meant to improve students' test scores fall prey of committing unintentional error of feminizing tutorial programs; that in turn cause fallback of male students' scores; that made them fail to benefit from their comparative strengths. Thus, it is commendable to take note of both male and female students' needs not only to empower and enable them benefit from their learning weakness but also from their strengths: gendering tutorial programs inclusive of male students.**

**Key word:** Tutorial, test scores, gender, Wollo University.

## INTRODUCTION

Gender gap in education has attracted huge body of educational research that examined the impact of sex and other social variables. Educational research on the gendered impact of tutorial scores is not as much as the huge body of research on the impact of general method of instruction, namely Single Sex Instruction (SSI) and Co-Sex Educational Instruction (CEI) on the pass rates, grades, and course choices of students in a SSI and CEI

universities and classes. These studies are cognizant of the widely held view that there is wide gender gap in pass rate, overall achievement, subject test score and professional choices in favor of their male counter parts. Women used to be considered underrepresented in higher education levels and achievements and subject specific scores the world over. However, recent studies have shown mixed results indicating lack of consistency

E-mail: [muuzgidey@gmail.com](mailto:muuzgidey@gmail.com) /[muuz.gidey@wu.edu.et](mailto:muuz.gidey@wu.edu.et).

Authors agree that this article remain permanently open access under the terms of the [Creative Commons Attribution License 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

in impact of type of instruction on gendered test scores, and the assumed degree of gender gap. A study that examined the effect of single-sex classes on the pass rates, grades, and course choices of students in a coeducational university found that one hour a week of single-sex education benefited females by 7% more likelihood to pass their first year courses and score 10% higher in their required second year classes than their peers attending coeducational classes. However, its effect on the probability of female students to take technical classes (they are underrepresented) and on male achievement is found to be nil. Furthermore, the potential explanation behind the single-sex positive effect for females was found out to be the effect of reduction in stereotype, threat against females; and it was not due to a potential tracking effect (Booth et al., 2013).

In the countries of the North, especially in USA, despite the ample improvement towards more women attendance of university education than men, SSI is sometimes promoted as a mechanism for narrowing gender gap. Behind this is the gender gap in standardized mathematics tests scores, especially at the top of the distribution. Proponents of SSI argued that females in SSI classes increases their test scores, especially in mathematics and science, and with greater likelihood of joining technical courses than their counterparts educated in mixed gender classes. However, this has been promoted against the dominant finding, as indicated above, the opposite (Booth et al., 2013; Pahlke et al., 2014). An experiment on the effect of Electronic Histological Tutorial (EHTI) on dental students test scores have found improvement but the statistical significance test proves the other wise; therefore, the finding was taken for inconclusive of the significance of instructional media in tutorial classes on student test scores (Rosenberg, 2006). Gender correlates were not included from the out set in the study, even though it were the over all finding has forestalled any further enquiry.

Studies incorporating both social and subject type variables were undertaken in an urban, at-risk high school primarily composed of individuals from disadvantaged populations to assess the effectiveness of single-sex instruction (SSI) on achievement outcomes, instructional practices, teacher efficacy, student behaviors, and classroom culture. The outcome of students grouped according to sex in Algebra and English classes were compared with coeducational students achievement results associated with SSI; the results were inconsistent, with gains shifting between groups. Differences in English achievement were not observed between SSI and coeducational groups. Standardized test results indicated superior performance for co-educational students. Important indication regarding

gender variation is that the SSI category was to provide a supportive environment for girls that accounts to inducement of greater participation and academic risk taking. Another finding related with this is that while teachers believed that the SSI environment was conducive to learning, students reject both the social and the academic benefits of SSI (Hoffman, 2008). This indicates the imperative for taking note of social dimensions outside of the class.

A study undertaken in SSI schools in English Speaking countries has found out the multiple difficulties to support the widely held view that SSI education benefits lower achiever boys, even more than female students in mathematics, physics and technical subject areas. The study has indicated the involvement of multiple factors in the findings of SSI-CEI comparative research outcomes precluding consensus around them. Among others, the national context of education gender is constructed in as SSI and CEI, the social environment of students, the admission system into SSI schools, the nature of subjects, the uncontrolled/controlled variables and methods researchers used have significant impacts (Smyth, 2010). A study undertaken in Korean middle school students to estimate the causal impact of school and classroom gender composition on achievements of students randomly assigned in to single-sex schools, coeducational (coed) schools with single-sex classes, and coed schools with mixed-gender classes have found significant impact of single-sex schools on male students' achievement; however, the impacts favoring SS student scores were not driven by classroom gender composition, but largely accounted for by increases in student effort and study-time. Therefore, the researchers concluded in finding little evidence that classroom or school gender composition affect the outcomes of female students (Lee, 2014). Despite the renewed interest in shifting from CEI to SS in the 21<sup>st</sup> century America, there is no consistent finding to support the claim that SSI do better than CEI rather than a mixed result in both directions (Gleason, 2011).

A critical examination of the SSI vs. CEI debate in USA, Rosemary C. Salomone gives a cautious balance and platform for Positive change. According to her there is "no clear indication that single-sex schooling harms students academically". In fact, SSI schools provide more positive attitudes toward certain types of male or female subjects in students of the opposite gender; for instance, disadvantaged minority students benefit both academically and socially from such schools. She accounts this to the wider access in SSI to leadership opportunities, the reduction of risk factors (e.g., teen pregnancy, drug abuse, etc.), and access to courses often gendered in coeducational schools emphasis of

single-sex schools (Salomone, 2003).

However, these opportunities do not necessarily presuppose associating them with SSI is always valid; inversely, it does not mean CEI is inherently unable to qualify them rather than the imperative for creating the environment in CEI schools. Two major points are crucial of Salomone to corroborate the above argument; first, the “effects of single-sex schooling and classes for boys and girls at the elementary, middle school, and high school level in rural, suburban, and rural contexts” are not yet examined; second, SSI can be good for some and “at its best, . . . an effective tool of empowerment and self-realization for some boys and girls,” and at its worst it can be “a tool of gender polarization and oppression”. Therefore, it remains to be how best to provide appropriate education for both sexes from different backgrounds at various stages of development than (Salomone, 2003) hard and fast commitment to SSI.

Nonetheless, other researchers challenge the above assertion based on research findings verging sexist orientation. According to Saguy et al., not from gender but biological point of view, the virtue of female SSI by male absence; they explain the highly visible male presence and dominance in class accompanied by female feeling of a male gaze at women’s body has two effects that reduces them to silence. One is what they called *objectification* effect of the male expectation to keep quite as an object of their attraction that makes them behave as per what is expected of them. The second is that the imposed passive dormancy generates anxiety about their external appearance that distract them from the task at hand. They warn against the subtlety of the process in real world that made it to be illusive for many researchers’ notice (Saguy et al., 2010).

Similarly, an in depth investigation of the effect of randomly assigned female students sitting next to each other in Chinese middle schools, unlike their counter parts considered in the same type of research in Korea, were found doing better if they sit near girls; sitting style was found to have no effect on boys achievement. Similarly, an extension of this research on the effect of being surrounded by male or female students on gendered student’s academic achievement found out that ‘. . . being surrounded by five females rather than five males increases a female’s test scores by 0.2–0.3 standard deviations but has no significant effects on a male’s test scores’ (Lu and Anderson, 2015).

The latter four researches, though did not confound there is a significant correlation between gendered educational achievement and SSI/CEI of students, in line with Salomone’s argument that SSI do no significant academic harm but may benefit some girls and some disadvantaged boys. However, the lack of consistency in

all countries and the potential danger of SSI to be “a tool of gender polarization and oppression “constituting the non-academic dimension indirectly affecting academic achievement is a dominant theme often over looked by proponents of SSI. The validity of the *objectification* and *sit next* arguments might have implications yet, do not seem to counter balance the far wider dearth of research finding. More so the erroneous assumption that opposite sex attraction does not apply to female students and those patriarchic roles are provided as reason to justify the superiority of SSI. This is like arguing for quarantine measures (appropriate for period of plague) for prevention of contagious disease for all times. Besides, these pro-SSI findings did not show female subject test scores in subjects like mathematics, physics and adoption of technical fields are narrowed.

On another platform, researchers like Hyde and Mertz have taken the task of challenging the intrinsic variability hypothesis; by examining math tests data in Minnesota and compared the numbers of boys and girls who scored in the top 5% of their year, they found narrow male – female variability for white Americans; girls higher than boys in the top 1% score. They also carried out test of Program for International Student Assessment (PISA) and the Trends in International Mathematics and Science Study (TIMSS) results in different countries with results supporting their claim; in Thailand, Iceland and UK 15 aged female greater than male; in similar results in mathematics scores in Netherland, otherwise equally competent. From these, they argued against biological variability by focusing to the non-applicability of hitherto assumptions to all populations regardless of ethnicity or nationality. Hence, gender difference in subject matter academic achievements were found out to be explainable by differences of students’ social background that encouraged and made large male achievement in those subjects possible. The greater number of male students in subjects like mathematics and physics as well as technical fields in turn exaggerates narrow gaps in academic achievement between male and female students to have big standard deviations. In effect, gender gap in such subjects and technical fields of appeared to be very wide and unbridgeable that it often is depicted as what it in actual fact is not. To substantiate their argument they invoke the absence of any finding in favor of the variability hypotheses in countries without gender gaps (Hyde and Mertz, 2009).

Recent findings in favor of non-variability hypotheses are currently multiplying; most are based on findings coming out from World first class universities in Countries of the North. A columnist at Discovery, ED Young, heralds similar findings in the Denmark, the Netherlands, the UK and Iceland, of equal girls and boys populating

the top 1% of the PISA results and gender equality in the top dozen countries (Yong, (a)2009; Yong, (b)2009; Yong, 2010). Therefore, they argued, wherever gender gap and achievement variations are maintained, the effects of stereotypes, a lack of female role models, willful misogyny and unconscious biases and hostile work environments are there (Hyde and Mertz, 2009).

Despite significant improvements in narrowing gender gap in Ethiopia, the hitherto discussed realities are farfetched and unheralded news. In short, female students are on the debt side of the balance sheet both in quality and quantity. More so does with the new third generation Universities, Wollo University is a member. The official avowal to gender mainstreaming and empowerment of women has often ended either in confirming 'female inferiority' or creating gender polarization often cause for masculinization of female specific needs. Otherwise, it means causing the disempowerment of their male counter parts. Wollo University gendered achievements have never been out of this scenario. So did students at College of Social Science and Humanities (CSSH), Department of Civics and Ethical Studies (DCESt). With the view to design far better inclusive and effective tutorial program in mind, the study, based on observations and trials carried out during 2013 and 2014 academic years to improve student test scores at DCESt, where the author is staff member has been undertaking experiments. Among efforts, how best should tutorials be designed and do tutorials have significant effects on student scores; if they do what of the gender dimension have been troubling the author. Accordingly, this study an outcome the experiment carried out to measure the effectiveness and the gendered impact of carefully devised tutorial program on student test scores.

### **Problem statement**

The investigator has been undergoing various efforts to improve student learning performance formally and informally in the department of CEST. Currently third year students have been in particular focused during 2013 and 2014 Academic years. Assessment of test scores in theory and practice of Peace and Conflict Resolution has shown steady improvement in test scores: male students outperforming the first one and half year. Following the launching of tutorial program thereafter, female students seemed to outperform their male counter parts, despite the male dominant learning spirit in classes. Nonetheless, except for mere assessment of improvement of test scores, the effectiveness, level of significance of tutorial classes, the differential effect on gendered test scores

and the level of significance thereof has not been established by research inquiry. Thus, the imperative to devise learning program for the improvement of overall gendered test scores student test scores on the subject required examining the impact of currently working tutorial practice. Therefore, the study is meant to fill the gap for the practical application of devising and launching effective tutorial program based on lessons from the findings of CEST test score associated with tutorial instruction.

### **Objectives**

#### **General objective**

To improve the effectiveness of tutorial classes on student test scores.

#### **Specific objectives**

1. To measure the significance of tutorial classes on students' Test scores.
2. To identify the impact of tutorial classes on gendered test scores of student.
3. To envisage effective and inclusive tutorial program on gendered students' test scores.

### **Hypotheses**

The objectives listed above were mixed, required both qualitative and quantitative inquiries. Accordingly, the two hypotheses herein below were designed to address the first two objectives.

As can be seen from the various studies discussed at length in the introduction there is no reliable consistent finding to assume tutorial classes do have significant effect on test scores, though improvement in scores are seen. Therefore, there is no evidence to support the tentative assumption that it significantly affects test scores. Hence, the first hypotheses rest on this as:

#### **Hypothesis 1**

$H_0$ : Tutorial classes have no significant effectiveness on students' test scores.

$H_a$ :  $\mu_1 = \mu_2$

$H_a$ : *Tutorial* classes have significant effectiveness on students' test scores.

$$H_a: \mu_1 < \mu_2$$

The second hypotheses are antecedent to the first, that since tentatively assumed have no significance, accordingly, the gendered test score's null hypotheses carries the same assumption as:

## Hypothesis 2

$H_o$ : Tutorial classes have no differential effect on students' gendered Test scores.

$$H_o: \mu_f = \mu_m$$

$H_a$ : Tutorial classes have significant differential effect on students' gendered Test scores.

$$H_a: \mu_f \neq \mu_m$$

## METHODOLOGY

### Epistemological claim

Epistemological claim is about how a body of knowledge can be generated from the subject of the study; as is detailed in the introduction there is no as such consensus around gendered student test score improvement and instruction; wider consideration of educational research also reinforces the contextual nature of results corroborating the former assertion. Therefore, since the immediate reality of solving educational problems should be the focus of educational research, the educational settings and problems can be studied using any method that accurately describes or solves a problem; that the research is based on pragmatic knowledge claim. In this spirit, the researcher is at liberty to strive to find ways to make education better that takes note of researcher active collaboration with participants to fully understand what works best in a given educational context. Accordingly, this study upholds to no theory or paradigm in particular emphasis other than based on its validity and relevance to the furtherance of success in achieving student test score improvement. Therefore, the use of theories and hypotheses is useful tools in helping to improve education were used from this consideration (Lodico et al., 2006; Cresswell, 2003; Dawson, 2007). Consequently, any special allusion to any theory or claim unless clearly stated or implied are to be considered out of the claim forwarded but the study.

In a more concrete manner, the objectives were meant to address issues requiring mixed type of inquiry that demands departure from the conventional commitment either to qualitative or quantitative strategy. Pragmatic orientation is an agile claim befitting the practical need of the study and the lack of consensus, implying context specific approach.

### Research design

The choice of design assumes the assumption set at epistemological claim that requires harmony all along the research process. Noting from the mixed objectives and the epistemological claim, the study has employed mixed design; this is not for

theoretical consistency rather than for actual operational reasons; for without the qualitative inquiry, the explanation for quantitative findings and inputs to address the third objective leaves unattended. The next question is how to blend the data; in accordance with the sequence of inquiry which begins by first establishing the significance of tutorial, followed by its gendered significance, then after the taking lessons on how to design tutorial classes for the same effect. Accordingly, based on pragmatic epistemological claim sequential explanatory design, the collection and analyses of quantitative data followed by the collection and analyses of qualitative data is applied (Babbie, 2010; Cresswell, 2003; Dawson, 2007). A before-After quantitative design based on the mean test scores of students is applied first and followed by qualitative explanation; the details on nature of data and techniques of collection and analyses are provided in the method below.

## Method

**Data source, type and tools:** The primary source of data is collected DCESt students and records, and notes taken from observations during class instruction in 2013/4, in teaching the course the theory and practice of Peace and Conflict. The rationale for selecting DCESt and the course deserves brief description. The author as student of peace and security has been teaching the course for four years at the DCESt and has control to both academic and administrative issues, and full account of the changes on students' performance every other year. Since, the course was not prerequisite, it is given only once to every batch of students. The course was selected because it has an elaborate maximum tutorial hour, six hours per semester determined by the national curriculum that makes it favorable to test significance.

The types of data were students' test scores, repeated evaluation, test and retest results of the semester rated from 50%, tutorial instruction considerations, classes' instruction student observation notes, staff performance reports and student-instructor face to face evaluation records as well as key student informant interview data were collected. To the qualitative sources semi-structured tools were employed; these were observation check list for document analyses and observation, and unstructured interview guideline. The choice of source, type and instrument of data collection were determined taking into consideration the practical financial, logistic and students attending the course in the academic year that, any question related with any of methodological considerations attributed to the nature of the study.

**Sample determination and sampling techniques:** the sample was drawn from the total number of students taking the course the 'Theory and Practice of Peace and Conflict' (3credit hour per week) in the academic year 2014 second semester; they 38 students constituting a section, but other student take the course. So the sample was taken from this population; the determination of tutorial hours bench marked the tutorial hour of the curriculum but trice, 18 hours, equivalent to six weeks regular class was organized as tutorial contact hours. Accordingly, the sample size was determined using the formula and table value developed by Krejcie and Morgan in 1970 provided below for the corresponding Chi-Square, degree of freedom and probability values (Sarantakos, 1993):

$$S = X^2 NP (1-P) / d^2 (N-1) + X^{2P(1-P)}$$

(Is given in the estimation table as 10) for a population (N) of 38 The TV of Chi-square ( $X^2$ )-1 degree of freedom at 3.841; Population

**Table 1.** Summary of data collection tools, sample size and sampling techniques.

| No | Population   | Sample size | Tool        | Sampling technique                                |
|----|--------------|-------------|-------------|---|
| 1  | 38students   | 10          | Test papers | Cluster random<br>$S=X2NP(1-P)/d2(N-1) +X2P(1-P)$ |
| 2  | 38students   | 6 (3f+3m)   | Interview   | Judgmental  |
| 3  | 18 T.Classes | 6           | Observation | Purposive hand picking                            |
|    | Total        | 22          |             |   |

Proportion of (P) at 0.05 or Confidence level of 95% and Degree of Accuracy (d) of 0.05: Sample size is 10.

The selection of individual students' test scores was carried out using random table taking every third score in the full list of 38 students in the class. Key informants were handpicked based on the researcher's judgment over the readiness and frankness of boldly telling the problems and successes of the tutorial program. Class observation with special focus on student activities was carried out during group activities (Table 1).

**Data Analyses and Validity**

Based on sequential explanatory analyses design, the analyses of quantitative data are given priority to test the hypotheses using the t-test statistical for small sample and normal population with unknown population variance. Test score mean and standard deviation along with their difference measures are employed with the significance level of the t-test as measure of validity. Qualitative data is used to substantiate qualitative findings using interpretive, analytic comparison, method of agreement and difference analyses. Therefore, finally quantitative and qualitative findings are one after another blended making the findings of the study.

**The significance of tutorial instruction on non-gendered and gendered test scores**

Significant test is the measure of how much statistical values do in actual fact show meaningful variability in terms of the variables considered. In this study, the variable under investigation, test score is tested in terms of time, i.e. non-gendered scores of students before and after tutorial; secondly, the same assessment is done in terms of gender variability. Therefore, in this section the test scores before and after tutorial are compared and the statistical significance of observed differences is check against the first null hypotheses under section 3.1 and the gender variability is checked in the subsequent section 3.2 against the second null hypotheses.

**Before-after tutorial analyses**

A random sample of two sample (before and after tutorial) test scores of ten students out of 38 student class size has been taken in order to know the effectiveness of tutorial attendance on student test scores. Successive evaluations out of 50 scores were administered before and after providing tutorial classes; the paired results are provided in a before-after fashion in Table 2. A before-After Tutorial Test Scores of students.

**Table 2.** A before-after tutorial test scores of students.

| Students | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 |
|----------|----|----|----|----|----|----|----|----|----|----|
| Before   | 21 | 31 | 43 | 25 | 29 | 31 | 40 | 30 | 44 | 23 |
| After    | 39 | 46 | 50 | 47 | 42 | 39 | 45 | 35 | 46 | 29 |

Before dealing with the gendered test scores and the differential effect of tutorial classes on male and female students to test the null hypothesis that  $H_o: \mu=\mu_{Ho}$  and the composite  $H_a: \mu\neq\mu_{Ho}$  it is proper to establish the effectiveness of tutorial classes in improving students test scores.

The hypothesis requiring test is that tutorial classes have no impact on mean score of students; in other words, the mean of difference between the two sample scores is zero:

$$H_o: \mu_1=\mu_2 \text{ which is equivalent to test } H_o: D=0$$

$$H_a: \mu_1<\mu_2 \text{ (to establish the tutorial has been effective)}$$

The appropriate hypothesis test for comparing two related samples involving small values of *n* that does not require the variance of the two populations to be equal but for the purpose of comparing are assumed equal and are expected to remain so is the Paired *t*-test provided by

$$t=D-0/\partial_{diff} / \sqrt{n} \dots \dots \dots (1)$$

Where

D=the average (Mean difference) of paired scores

$\partial_{diff}$ =Standard deviation of differences

n=Number of matched pairs (Cothari, 2004)

$$D=\sum D_i/n \dots \dots \dots (2)$$

Accordingly, the value of variance of difference is provided by (the function of  $\sum D_i^2, D^2$  and *n*)

$$(\partial_{diff})^2=\sum D_i^2- (D)^2.n/n-1 \dots \dots (3)$$

Therefore, based on the values of formula 2 and 3 above, and taking note of the normal distribution and independence of differences the significance of mean difference can be judged by computing the values of the test statistic *t* of formula 1 fixing the rejection region at 5 percent level using one-tailed test (in the left tail because  $H_a: \mu_1<\mu_2$ ) and degrees of freedom at (n-1) (Cothari, 2004).

Hence, the mean and standard deviation of difference of the two matched samples denoted by  $X_i$  and  $Y_i$  respectively is shown here in Table 3.

**Table 4.** Paired samples of student test scores: before and after tutorial.

| Students | Sex       | Score Before Tutorial<br>$X_i$ | Score After Tutorial<br>$Y_i$ | Difference<br>( $D_i = X_i - Y_i$ ) | Difference Squared $D_i^2$ |
|----------|-----------|--------------------------------|-------------------------------|-------------------------------------|----------------------------|
| 1        | F         | 21                             | 39                            | 18                                  | 324                        |
| 2        | F         | 31                             | 46                            | 15                                  | 225                        |
| 3        | F         | 43                             | 50                            | 7                                   | 49                         |
| 4        | F         | 25                             | 47                            | 22                                  | 484                        |
| 5        | F         | 29                             | 42                            | 13                                  | 169                        |
| 6        | M         | 31                             | 39                            | 8                                   | 64                         |
| 7        | M         | 40                             | 45                            | 5                                   | 25                         |
| 8        | M         | 30                             | 35                            | 5                                   | 25                         |
| 9        | M         | 44                             | 46                            | 2                                   | 4                          |
| 10       | (F=5;M=5) | 23                             | 29                            | 6                                   | 36                         |
| n=10     |           |                                |                               | $\sum D_i=101$                      | $\sum D_i^2=1405$          |

Mean of difference is,

$$D = \sum D_i / n = 101 / 1405$$

$$D = 0.072$$

Standard deviation of difference is

$$(\sigma_{diff})^2 = \sum D_i^2 - (D)^2 \cdot n / n - 1$$

$$(\sigma_{diff})^2 = 1405 - (0.072)^2 \cdot 10 / 10 - 1$$

$$(\sigma_{diff})^2 = 1405 - 0.052 / 9$$

$$\sigma_{diff} = \sqrt{156} = 12.5$$

Hence,

$$t = D - 0 / \sigma_{diff} / \sqrt{n}$$

$$t = 0.072 - 0 / 12.5 \sqrt{10} = 0.072 / 3.94 = 0.018$$

$$\text{The degree of freedom} = n - 1 = 10 - 1 = 9;$$

To determine the rejection region of the alternative hypothesis at 5 percent risk comes to as under, using the table of *t*-distribution table for 9 degrees of freedom:

$$R: t < 1.833 \text{ (Cothari, 2004)}$$

The observed value of *t* is 0.0018 (from the R:  $0.0018 < 1.833$ ) falls within the rejection region that  $H_0$  is rejected and conclude difference in test score before and after tutorial is too significant to accept the null hypothesis and assume the difference was due to sample fluctuation. Therefore, tutorial attendance was effective on test score improvement. A-test result similarly has confirmed the same significance at the same degree of freedom and 5 percent level of significance. Literally, it means there is enough reason to accept the assumption that tutorial classes have significantly affected positively the overall score of the sample student scores.

However, the variability between male and female students given their respective before tutorial scores cannot be inferred from this. Therefore, to test the second hypotheses the gendered variability is presented in the next section.

### Testing the gendered significance of test scores

From the above statistical evidence as there is no reason to accept the  $H_0$  and Reject the  $H_a$  implying strong significance of tutorial classes on test scores; nonetheless, it does not show as to whether

it has gender differentials of test scores. To restate the initial hypothesis:

$$H_0: \mu_f = \mu_m$$

and

$$H_a: \mu_f \neq \mu_m$$

Taking the same considerations for the test of significance of effectiveness of tutorials utilized above and the sample scores in Table 1, we need to reconstruct gendered after tutorial scores to calculate mean and variance of female and male students (Table 4).

Since the population variance is not known and the sample size is small, hence the *t*-test for difference in mean, assuming the population to be normal and the test statistics *t* can be computed as (Table 5),

$$t = \frac{X_f - X_m}{\sqrt{(n_f - 1)\sigma_f^2 + (n_m - 1)\sigma_m^2}} \cdot \sqrt{\frac{n_f + n_m - 2}{n_f + 1/n_m}} \text{ (Cothari, 2004)}$$

Substituting the table values in the above formula we get.

$$t = 21.3 \quad 8.6$$

With d. f.  $= (n_f + n_m - 2) = 5 + 5 - 2 = 8$  for two tail test at 5 percent significance level, the test statistics table value of *t* is 2.306. The rejection region of  $H_0$ ,  $H_0: \mu_f = \mu_m$ , is fixed at

$$R: |t| > 2.306$$

The observed value of *t* 21.3 falls within the rejection region that, there is no statistical evidence to reject the alternative hypothesis,  $H_a: \mu_f \neq \mu_m$ , that tutorial classes have had differential impact on gendered scores in favor of female students. However, the hypothesis test does not tell what other alternative explanations are behind far better test scores of female students after attending tutorial classes. Yet, it is obvious from the finding that male test scores have become on the debt side of the balance sheet, asserting the absence of evidence to accept the null hypotheses that tutorials have non-differential impact on gendered test scores. Inversely, it means there is enough evidence to accept the assumption provided by the alternative hypotheses that the variability has less positive significance to male than to female

**Table 5.** Gendered After tutorial test score mean and variance.

| Female students' sample scores |                |                   | Male students' sample scores |                |                   |
|--------------------------------|----------------|-------------------|------------------------------|----------------|-------------------|
| Score after tutorial           | Difference     | DifferenceSquared | Score after tutorial         | Difference     | DifferenceSquared |
| $X_{fi}$                       | $(X_{fi}-X_f)$ | $(X_{fi}-X_f)^2$  | $X_{mi}$                     | $(X_{mi}-X_m)$ | $(X_{mi}-X_m)^2$  |
| 39                             | -5.8           | 31.9              | 39                           | 0.2            | 0.4               |
| 46                             | 1.2            | 1.4               | 45                           | 6.2            | 38.4              |
| 50                             | 5.2            | 27                | 35                           | -3.8           | 14.4              |
| 47                             | 2.2            | 4.8               | 46                           | 7.2            | 51.8              |
| 42                             | -2.8           | 7.8               | 29                           | -9.8           | 96                |

$$\sum X_{fi}=224; \sum (X_{fi}-X_f)^2 =72.9; \sum X_{mi}=194; \sum (X_{mi}-X_m)^2 =201; n_f=5; n_m=5.$$

**Table 6.** Summary of sample mean and variance.

| Sample group | Mean scores | Standard deviation | Sample size |
|--------------|-------------|--------------------|-------------|
| Female       | 44.8        | 4.3                | 5           |
| Male         | 38.8        | 7                  | 5           |

student test scores. As to whether this means female students have outperformed their male counter parts or there was some other explanation appears requiring further inquiry that corresponds along with the third objective of the study. Therefore under the next subsection qualitative inquiry is presented to find the missing link; but before that the implications got to be briefly discussed to set the ground for the subsequent discussion.

### IMPLICATIONS OF FINDINGS: BEYOND THE STATISTICAL MEASURE

Studies on gendered test scores collaborating significant achievement or subject score improvements benefiting female students are available based on the modality of instruction: SSI vs.CEI; but particularly with tutorials. Instead, to make logical inference a tutorial program that benefited lower achiever student group was not expected to worsen off before tutorial better of group: female and male students respectively. To compare the finding of this study to the research findings showing female students benefiting from an instruction could be one alternative to probe the direction of subsequent inquiry. However, unlike the finding of the study, male students being worsen off, male students are either unaffected as in the case of the *sit next* experiment Korean (Lu and Anderson, 2015) or only disadvantaged boys were found benefiting from SSI tests in USA school system (Salomone, 2003). So, what this implies to the nature of tutorial requires explanation.

Moreover, another issue requiring further investigation and proper explanation is the comparatively worsen off

status of male scores who had been better off before tutorial. While preceding research findings do not provide supporting the finding of the study, it is hardly possible to take it unexplained. On the other hand, the only available statistical explanation is unable to provide answers to this paradox but is sufficient evidence to embark further qualitative inquiry complimenting the above finding. In order to address this paradox, the study embarked on qualitative inquiry based on the information from observation notes of male and female students during the process and further interview with purposively selected student informants were undertaken providing explanation under the next section. This section addresses how best to design tutorial class for the betterment of both sexes as set out in the third specific objective in the introductory part of the study.

### The missing link on the *feminization* of gendered test scores

The main objective of the tutorial program was to improve the test scores of both female and male students of the class; lower performing female students and higher performing male students are assumed to show Positive significant effects. Otherwise, other things, *citrus paribus*, one would not expect higher performing male students would be worsen off, at least. Yet, the finding shows the contrary that inquiry in to the overall context of the tutorial and the process got to be examined for any explanation to the anomaly observed. Therefore, the following sub-



sections are devoted to this utility. The *Feminization of Gendered Test Scores* used in the title of the section is meant to connote the unexplained condition of favoring female students but not their male counter parts.

### Setting the context of tutorial delivery

The learned experience from preceding lunch break and late afternoon classes as well as make up classes have shown student drowsiness, fatigue and disorientation that accounts to mostly long class attendance and discomfort with loosing vital week end time to attending classes. Student-instructor joint evaluation and evaluation report of the investigator exposed that students rate him as the best but complain about long hour teaching, make ups during week end days that goes against their interest to use it for social and religious engagements (DCESt-1, June 2014) (DCESt-2, June 2014). Moreover, the investigator observed students have silent complain on not regularly getting involved in party meetings during class hours.

Launching the tutorial classes has taken serious note of the counter-productive effect of these contests and the imperative to avoid them. Accordingly, tutorial hour was fixed to one hour regular attendance on Saturdays in the morning after church attendance that begins at 3: 00 local time with the full consent and free will of students. No obligation was attached to attendance and student feedbacks confirmed the same. At times, students took the advantage of convening classes favorable to all students; classes had been arranged accordingly to which positive feedbacks were offered. Therefore, the context was set in free and participatory manner in line with avoiding setbacks to the learning process. Thus, the context provides no explanation to the paradox.

### Process analyses of tutorial delivery

During the process the behaviors of male and female students in terms of class participation and motivation have differential expression; mostly male students despite the effort to revitalize the active spirit they usually show at regular classes are seen in disinterested mood. On the contrary female students had been all time all excited exhibiting unflinching and whole hearted interest. Motivation to ask questions and respond to provocative discussions, when even left alone inviting male student participation female students were taking the chance to reflect. It seemed that they were availing of the venue to compensate the feeling of being disfranchised in regular classes by male dominance. Female informant views confirm the same:

*“There are things troubling to female students which I cannot tell you in relating with male friends; to ensure survival there are many things demanded from girls; tutorial classes require nothing in return except attendance; there is also wide space for female voices because male students resign from participating by looking down at tutorial classes as the class for the weak and female students. (Student-F1, 2014)”*

Issues of sexual manipulation and feeling of dependency on male colleagues and the peer influence of being judged loser are highlighted as motivation to work as hard as possible in these classes. The view of male students is at total variance with the above view except the last attitudinal factor.

*“Mostly female students except attending Sunday prayers they do not have as much social engagements as male ones; classes on Saturday are not moody mostly though we like the delivery and the subject matter. Also, up to now attending tutorial classes are viewed as that of the meek minded and girls those most male students do not take it seriously. After all, had we been good at our study regular classes are more than enough to score good grades” (Student-MI, 2014).*

The above view is indicative of how much the prevailing attitude in the wider campus environment and the age old patriarchic belief in female inferiority and down cast evaluation of the other may have contributed to the male comparative decline. The issue of social engagements and feeling of disfavor Saturday classes for male is only mere shame for it was consensually accepted by all students.

According to the above female informant test score improvements of female students at times exceeding high scorer male class mates have boost their confidence and that in turn decreased male-female usual relations and cooperation. In her words,

*“I do not mean all boys are ill motive. There are many good hearted class mates who treated us and even help us in an informal ways; yet, among us it could be wrong the widely held view is that, boys are not happy with the fact of girls independently working and achieving more than boys (is considered illogical)” (Student-MI, 2014; Student-FI2, 2014)*

She goes on to touch another development,

*“I cannot hide from you that with encouragement and change our relations with the boys have been changing from our side. I think, perhaps, some girls might even have unpleasant memories . . . of remorse on how they used to give in to their demands for the mere opportunity to be included in achiever male groups” (Student-FI2, 2014)*

Formal and informal observations of relations in class and out of the class do support the above reflection though it still carries the “we” and “they” unfriendly category. However, important point to capitalize is the powerful of quality and content of social interaction and interdependence, or in general relational considerations for the inclusive success of a given instructional and learning program which was not considered in the process. Prevailing positive relational assets before any program has to be put to the utility of furthering social and relational solidarity crossing beyond whatsoever social divides. Moreover, the underlined statements are also indicative of the need for capturing subjective views and possible future scenarios to avoid undesired results from well intentioned efforts.

Another important factor is pointed out by a couple of female and male participants. For both female and male students, even though the program was initially given to them as part of regular process in the course out line and oriented accordingly, it was taken for the instructor’s self motivation to empower female students that subtly pushed male students to self resignation. The intonation of female articulation that, *“We know how much you work to help us how could we let you down! We always talk among us that we should do instructor proud. Male students very well know and make fun of us; yet, we did it in action!”* (Student-FI5, 2014), is much telling to what direction the change swung relationships. The sense of the defiance and jubilation is pushy on the other side.

The level of concern and diligence of the instructor and perhaps the care exhibited to encourage female students might have been formed the image that it was meant to female class mates and was personal mission to help them which the instructor did not set control of. The suggestion from male students’ world by word reiterated the same.

*“Of course we were making fun of them; we were saying ‘it is not by the number of ours class attended but effectiveness!’, even though we like very much you classes; it seemed as though it was designed for female student consumption. They also talk much about these classes and you as making favor for them”* (Student-MI2, 2014).

Another male informant see the problem from departed vintage point saying that “since we have every necessary material provided to us and believed it to be mere repetition most male students were lenient” (Student-MI3, 2014). On the other hand, a female informant recants that

*“. . . among the boys in addition to the negative attitude and chauvinism, they used to judge your efforts as undeserved excessive care to girls and even very few used to say it was mere quest for cheap popularity in your part. Even while seeing far improving female test*

*scores they were resistant to accept the fact. It was totally their problem”* (Student-FI6, 2014).

From the above point we can see how the dominant trend in campus comes to define good intentioned efforts as “. . . cheap popularity!” that becomes counter-productive to the very effort as seen by male students attitudes. The investigator’s post facto recollection and reexamination perhaps partly indicates the above view and exposes the unintended but silent assumption that ‘a method of teaching that can empower less performing students (most of them female) cannot not make well performing parties even by far better off than ever; the slight possibility of such errors might have unintended consequences of Feminization of efforts. Hence, instructors are supposed to make ginger mix of empathy, care, diligence and dispassionateness to achieve desired improvements in their students.

Add to these, the excitement of instructors to see hitherto less performing students with new energy and courage, as the experience of the investigator educator had been, have the power of taking away timely and critical examination and rectification of the process slipping in to the undesired.

## DISCUSSION

The blended findings of quantitative and qualitative inquiry discussed in the preceding sections are instructive of making serious reflection relating with the findings of other research works discussed in the introduction. As apart from the SSI vs.CEI debate, important lesson to take pertinent to this study is that lower performing female students provided they are given the chance to attend tutorial classes not dominated by male leadership, participation and attention may perform significant test scores. A point to ponder is that despite many findings confounding the absence of effects on male subject test scores and overall educational achievements in the finding of this study shows the otherwise.

This in turn takes the focus to another important theme raised by Salomone about the chance of benefiting female students and some disadvantage boys but without causing significant harm to no one. However, the findings of this study do contradict this assumption on the ground that, as can be seen from the significant variability in favor of female students and from the male informant’s complaints quoted in the preceding section. This accounts to the absence of gender stereotypes against female students by the careful handling of the tutorial that, unfortunately, was accompanied by the activation of social stereotype dominant in the University campus. Tutorial classes are often perceived at various levels of

Wollo University academic leadership mechanisms of reducing dismissal rate. Hence, are often taken programs for the under performer than the medium and high achiever. As indicated at the beginning, the latter category was dominantly male constituency which resulted in male students being stereotyped.

Besides, male students' perception that instructors were carrying out tutorial classes to the advantage of female students considered under performer adds to the furtherance of negative energy described by Salomone as the formation of ". . . *gender polarization and oppression*". In this case, the effect of the polarization and oppression has taken the form of deterioration of male-female social relation articulated by male students as instructors' favoritism towards and to gain cheap popularity from female students. On the other hand, female students viewed the unfavorable male response as the unhappiness to see them independent of male support and the untold manipulations accompanying it. Though the objectification argument by no means is fully acceptable for reasons discussed in section one, however, observation of tutorial classes and the subtle indication connoted by the term '*untold manipulations*' could be taken unintended externalities that follow with gender polarization and degeneration of healthy relationships in a CEI University like Wollo.

Although the study was carried out in CEI University, in terms of the impact of excessive encouragement and support to male and female students, it corresponded with findings of SSI critiques that one sex would be disadvantaged; in this case male students being relatively disadvantaged. However, unlike SSI critiques, this was not because of the problem of SSI in selective criteria of interested and better students as factor for good performance; because in this case, tutorial classes created opportunities for active participation of all students but female students availed themselves of the opportunity more than male students; male prejudice against tutorial classes as designed to help the weak and female was a factor that precluded them from benefiting from tutorial classes as significant as female students.

However, vital point not found in all of the findings discussed at the introduction is the ample possibility of causing significant impact; while preceding studies focused on narrowing the academic performance variability and overall gender gap, this study has engaged in testing how to significantly improve test scores of. Most studies from SSI schools or class rooms have shown contrary findings to the idea of causing significant improvement of female student test scores. In polar opposition to these findings, this study set in CEI University and mixed class has indicated the importance of paying sufficient attention to preclude value reversal,

which this author termed as the *Feminization* of tutorial classes.

Another important point departing from above discussed findings is the possibility of impacting significant change in student test scores without paying a heed to contemporary global debate for and against SSI or CEI.

The overall theme of vital concern is that, in launching any learning program it is vital to take note of factors seemingly far and independent of the process; how good intentions and well though procedures differently affect the different groups is determinant of the strength of the light at the end of the tunnel; for some too bright to see, dazzling and sometimes blinding. Comprehensive considerations of psychological and relational dimensions of learning are to be cared for.

## CONCLUSION AND RECOMMENDATIONS

The summary conclusion and recommendation of this short examination may have far reaching implications for wider study, nonetheless, for the purpose of this study it is reduced to the following major conclusions and corresponding recommendations.

### Conclusion

Tutorial classes in CEI University, even as imperfectly as the tutorial in this study, can and do significantly affect student test scores; more so, greater significant test scores can be achieved for the lower achiever female students. The success in this respect is important contribution for overall improvement of academic achievement in CEI tutorial classes. Because, succeeding in improving the least performing student to benefit, provided externalities are controlled from disturbing the potential of the high achiever to be distracted, can be an indicator that teacher educators with sufficient devotion and effort can design tutorial programs for greater success. The findings of this study can be taken as one piece of proof against the argument for SSI to improve student test scores.

Excessive emphasis on realizing test score improvement of lower achieving students may overshadow basic social and inter-personal relations between male and female students that may lead in to gender polarization and oppression. This, in effect, could be a case for the imperative for not only enabling all to benefit from their strengths but also precluding the possibility of being worsen off from their weaknesses.

The success of tutorial classes in affecting significantly test scores are negatively affected by extraneous

variables, social stereotypes, and gender polarization; not the least, the Positive externalities in the form of the chance for participation and leadership as well as sufficient support and encouragements behind significant improvement in test scores.

The worsen off status of male students is instructive of the imperative to take care of disempowering better performing male students by the empowerment of female students measured in terms of test score relative betterment; this presupposes the practical possibility of realizing both.

Therefore, experience shows the need for painstaking perseverance of multiple actors in the process of learning; success and failure are camouflaged in each others' skin impeding clear vision. The very scanty analyses of test score based analyses shows success at improvement in test results but also failure in basic inter-personal relations vital to pedagogy of active and independent learning. This has emanated from both personal assumptions, the wider social, cultural, relational and inter-group dynamics in our case gendered relations which are not garnered for the utility of far greater success.

## RECOMMENDATIONS

Therefore, it is commendable to work on the following points in devising teaching and learning programs in general and tutorial programs in particular.

First, further study on gendered test score analyses in the University to confound further reflection for practical utility of significantly improving test scores of both sex got to be carried out in the following year. In so doing, the following points should be considered as tentative recommendations.

Second, tutorial classes should design mechanism of reducing gender polarization and repression should be carefully taken care of;

Third, the design and implementation of tutorial classes (however every well intentioned) should be based on critical assessment of the objective and subjective interpretation of gendered attitudes, perceptions, vulnerabilities and feeling of threats; these matters got to be exhaustively dealt with before launching the program and meticulous supervision should be done during the implementation;

Fourth, micro efforts at departmental and college levels, like this study, have to take note of envision fixing problems associated with the wider University community upheld perception, attitude and response to greater success; so that negative attitudes upheld by educational quality organs to tutorial classes as dismissal reduction

programs and their impact on student prejudice on tutorial classes have to be addressed. Hence, future tutorial programs should be Janus faced: turned to the past, observant of the present and aspirant of the future practices and value systems.

Fifth, in view of the third recommendation, the need to take the big picture of university society and national education system should not be taken for extremely exaggerated and out of reach point for it brings to focus the interplay of factors provided preemptively taken note of could be turned in to opportunities;

Sixth, every educational program should be molded with cursory attention to avoid both the Feminization and masculinization of the final outcome as well as every bit of the process leading to the imagined improvement.

Seventh, modalities of transforming emergent results of success and the earlier better capabilities in functional integration with the emergent has to be articulated before embarking new programs.

Finally, to avoid the potential of abuse by students, the design and implementation of tutorial classes should be impersonalized as official and regular duties of instructors than expression of unique commitment of individual educators.

## Conflict of Interests

The authors have not declared any conflict of interests.

## REFERENCES

- Booth AL, Cardona-Sosa L, Nolen P (2013). Do Single-Sex Classes Affect Exam Scores? An Experiment in a Coeducational University. Discussion Paper No. 7207. Germany: Bonn: IZA.
- Booth AL, Cardona-Sosa L, Nolen P (2013). Do Single-Sex Classes Affect Achievement? An Experiment in a Coeducational University. *J. Educ. Res.* 1(2).
- Babbie E (2010). *The practice of social research: 12th edn.* Belmont, CA: Wadsworth.
- Hoffman BA (2008). The Effect of Single-Sex Instruction in a Large, Urban, At-Risk High School. *J. Educ. Res.* 102(1):1-50.
- Cothari C (2004). *Research Methodology.* New Delhi, India: New age international (p) limited, publishers.
- Cresswell J (2003). *Research designs: qualitative, quantitative and mixed methods approaches, 2nd ed.* Thousand Oakes, CA: Sage Publications.
- Dawson C (2007). *A practical guide to research methods 3rd ed.* Oxford: How to books Ltd.
- Pahlke E, Janet Shibley H, Carlie MA (2014). The Effects of Single-Sex Compared With Coeducational Schooling on Students' Performance and Attitudes: A Meta-Analysis. *Psychological Bull.* 140(4):1042-1072.
- Lu F, Anderson ML (January 2015). Peer Effects in Microenvironments: The Benefits of Homogeneous Classroom Groups. *J. Labor Econ.* 33(1):91-122.
- Gleason P (2011). *Gender Segregated Learning Environments An Analysis Of The Perceived Impact Of Single-Sex Classrooms In South Carolina.* np: University of South Carolina.

- Hyde J, Mertz J (2009). Gender, culture, and mathematics performance. *Proc. Natl. Acad. Sci.* 106(22):8801-8807.
- Lee S (2014). All or Nothing? The Impact of School and Classroom Gender. *J. Educ. Learn. Code: I20, I21(24):1-50.*
- Lodico Dean T, Marguerite G, Spaulding, Katherine H Voegtle (2006:11). *Methods in Educational Research From Theory to Practice.* San Francisco, USA: Jossey-Bass A Wiley Imprint.
- Rosemary C, Salomone A (2003). *Same, Different, Equal: Rethinking Single-Sex Schooling* New Haven, CT, No. 304 . Harvard Educational Review. New Haven : Yale University Press.
- Rosenberg H ( December 2006). Effectiveness of an Electronic Histology Tutorial for First-Year Dental Students and Improvement in "Normalized" Test Scores. *J. Dental Educ.* pp.1339-1345.
- Saguy T, Quinn D, Dovidio J, Pratto F (2010). Interacting Like a Body: Objectification Can Lead Women to Narrow Their Presence in Social Interactions. *Psychological Science.*
- Sarantakos S (1993). *Social Research.* Basingstok : Macmillan.
- Smyth E (2010 ). single-sex education: What does Research Tell us? *Revue française de pédagogie* 171 avril-mai-juin, 47-55.
- Yong E (2010, January 12, 9:30 am). Discover: Science for the Curious: How objectification silences women – the male glance as a psychological muzzle . Retrieved Januar 10, 2015, from [www.DiscoverMagazine.com](http://www.DiscoverMagazine.com) .
- Yong E (2009a). June 1, 4:51 pm.). Discover: Science for the Curious: Gender gap in maths driven by social factors, not biological differences. Retrieved January 10, 2015, from [www.DiscoverMagazine.com](http://www.DiscoverMagazine.com).
- Yong E (2009b). June 23, 12:00 pm). Discover: Science for the Curious:Hidden beliefs in science stereotypes predict size of gender gap across 34 countries. Retrieved January 10, 2015, from [www.DiscoverMagazine.com](http://www.DiscoverMagazine.com).

## CITATIONS

- DCESt-1. (June 2014). Staff Performance Evaluation Annual Report: 2013/4. Wollo University, CSSH, DCESt: np: DCESt.
- DCESt-2. (June 2014). Face to face student-instructor joint evaluation held on Jun, 2014. Wollo University, CSSH, DCESt: np: DCESt Archive.
- Student-FI (2014, September 04 ). On the Purpose and Quality of Tutorial Classes. (G. Mu'uz, Interviewer).
- Student-FI2 (2014, September 9). On the male responses on the Positive Impact of Tutorial Classes of female student scores . (G. Mu'uz, Interviewer).
- Student-FI5. (2014, September 10). The Impact of Tutorial on female student test score and male student Relations . (G. Mu'uz, Interviewer).
- Student-FI6. (2014, September 12). On Post-Tutorial Social Relations . (G. Mu'uz, Interviewer).
- Student-MI. (2014, September 8). On the Impact of Tutorial Classes . (G. Mu'uz, Interviewer).
- Student-MI2. (2014, September 11). Evaluation of Tutorial cassess and Instruction . (G. Mu'uz, Interviewer).
- Student-MI3. (2014, Septemeber 12). On the Importanc e of Tutorial classes . (G. Mu'uz, Interviewer).