INTRODUCTION

Many countries are concerned about the low number of women in science (D’Andola, 2016). Many researchers have investigated possible causes for why there is such a low number of women in sciences. For example, Moss-Racusin et al. (2012) found that there was a subtle gender bias favoring male undergraduate students in their double-blind study. Moss-Racusin et al. (2012) reported that male applicants were rated as more competent than their female counterparts for laboratory manager positions from faculty members of biology, chemistry, and physics departments. The role of nurture has been shown to be an important factor in the gender gap in spatial abilities by Hoffman et al. (2011) in their study of eight tribes in India. Harsh et al. (2012) demonstrated, however, that participation in an undergraduate research program before graduate school was effective for males and females in contributing to the pursuit postgraduate participation in chemistry and physics.

A search of the literature highlights that there is literature concerned with gender differences in physics and mathematics, but few studies have been dedicated to the field of chemistry, especially chemistry taught at the university level. Two of these studies were Shibley et al. (2003) which examined cognitive ability in chemistry and Turner and Lindsay’s (2003) on cognition and organic chemistry. This paper study sought to examine university level chemistry, specifically to highlight if organic chemistry and physical chemistry are actually harder for one gender over the other. To address this question, the collected final scores of two modules of study over a 15 years’ period were analyzed. This analysis was undertaken to identify whether teaching with a gender perspective, an emphasis on reasoning and construction of knowledge, using molecular models in taught sessions, and applying open-book examinations had effect on the scores of students. Particularly, on whether there exists an interaction effect between student’s scores and gender.

The university autonomous metropolitan (UAM) was founded in 1974 and was the third public university in Mexico City. The student population is about 42,000 students, which is distributed over five campuses. In four of these campuses, the traditional teaching methodology was followed; however, the Xochimilco Campus applied a module teaching method. The goal of this module teaching method was that the students would engage in active learning and have the ability to gather and analyze information. The assessment of the learning involved a range of artifacts such as homework, seminars, and included written examinations, in which the close-book examination method has been preferred by many teachers in Mexico.
The Bachelor of Biological Pharmaceutical Chemistry (QFB) offered by the UAM Campus Xochimilco (UAMX) has a career scheme oriented to the pharmaceutical industry and comprises twelve modules divided up as three modules per year for the 4 years of study. Each module requires a total of 312 h of both classroom and laboratory participations. These hours consist of 3 days of 6 h per day of classroom theoretical work and 2 days for practical work per week. Two of the twelve modules are organic chemistry of drugs (O) and drug production (P), which are taken at the beginning and at the end of the student’s 2nd year of study, respectively. Module O concerns organic chemistry, and module P has to do with physical chemistry on basic unit operations and includes organic chemistry.

From 1974 to 2010, the enrolment of women in the QFB degree has increased almost 8 times compared with that of the men’s 5-fold increase (Annual Report UAM, 1974; 1980; 1990; 2000; 2010) (Figure 1).

However, only about 50% of the women and between 35% and 48% of the men completed the QFB bachelor degree. This completion rate corresponds with the UNESCO Institute for Statistics (2017) about the general data for bachelor degrees in 124 countries. However, it would be important to find out how to increase these completion percentages.

Worryingly by 2010 not only did the percentage of graduation in the QFB bachelor degree fall by about 14% but also the career earnings of female graduates of the degree was half that of male graduates (Annual Report UAM, 1974; 1980; 1990; 2000; 2010) (Figure 2).

**MATERIALS AND METHODS**

The participants of this study were students in the 2nd year of the QFB degree at the UAMX. In their 2nd year of study, the student’s average age was 20 and the majority self-identified as belonging to the middle and low middle classes. The majority of the participants were born in Mexico. None showed any visible disability nor declared to have one.

The participants were the 343 undergraduate students during 1995–2010. The data collected were analyzed in two subgroups: 166 were enrolled in the module O and 177 in the module P. The O subgroup had 100 women and 66 men; meanwhile, the P subgroup had 115 women and 62 men.

The module organic chemistry of drugs (O) dealt with organic chemistry. In this module, students made use of analyses, synthesis, as well as spatial skills (i.e., in stereochemistry, absolute configuration and concerted reactions). This module did require some rote memory skill. The module drug production (P) involved physical chemistry for basic unit operations of mass and energy balance. Mathematics, physics, and chemistry were used for resolution of the problems related with the unit operation in the pharmaceutical industry.

As stated, modules O and P were placed at the beginning and the end of the 2nd year of the QFB bachelor, respectively. The modules were divided into 18 h of theory and 12 h of laboratory per week for 3 months. The content of the theory was presented by the same professor and with short seminars led by the students. The final scores were the average of the evaluations obtained from the seminars, participation in classroom, homework, written assignments, examinations, and experimental work. The experimental work was in all cases passed with the written experimental report, which was carried out by teams of students.

The final scores ranged from 0 to 10. Students were assigned a final score with letters, in which MB meant that the student obtained a very good score between 8.334 and 10, B (good score) 7.334 and 8.33, S (sufficient) 6.0 and 7.33, and a failing score of NA was between 0 and 5.9.

There were four examinations for each module. The examinations were written by the same professor who taught the modules. The examinations comprised of five problems or questions which required higher order thinking or analysis of new situations to include the student’s written justification. Each examination lasted a maximum of 3 h, and each problem was scored. The students were allowed to use a calculator, their notebooks, and books.

Statistical analyses of the results were computed separately for females and males for each module with Statistical Package...
for the Social Science, and data were presented with raw final scores and percentages; \( \chi^2 \) test with a statistical significance 0.05 (\( \rho \)) and Yates correction (Hatch and Lazaraton, 1991) were used to determine if there was association between the study’s variables: Gender and scores. For data that included less than \( n < 5 \), to prove if there was any association between the higher scores and the groups of the same module, the \( \chi^2 \) was applied without Yates correction with a significance level of 0.05 (\( \rho \)).

**RESULTS**

Of the total participants (\( n = 343 \)), 63% (\( n = 215 \)) were women and 37% (\( n = 128 \)) were men. The samples for module O were 166 (48%) and for module P were 177 (52%). The population of women as participants was the majority with 60% and 65% for module O and P, respectively. The higher female/male ratio did not affect the data because gender analysis of the data was conducted wherever appropriate.

In terms of score distribution among modules O and P, it was observed that there was a difference within the proportion of high scores without taking gender into account, as shown in Table 1.

Even though the O module had fewer students than the P module, the percentage of the lowest scores was 54% and 57%, respectively.

Tables 2 and 3 show the variation in gender in the O and P modules. More female students had higher MB and B scores than their male counterparts in the O module and higher MB scores in the P module.

Association between the scores and gender was found for the O subgroup (\( \rho = 0.04 \)) but not for P subgroup (\( \rho = 0.26 \)) (Tables 2 and 3). No statistical differences were found in the proportion of those students who could not pass each module (\( \rho = 0.52 \)).

Two analyses were made to avoid bias on the association between the type of module and the scores needed to pass the module. Table 4 shows the number of students with MB and B scores and the number of students with failing scores into both modules.

Meanwhile, Table 5 shows the number of students which passed the module with an S score compared with the failing NA score in both modules. Neither of the analyses showed any association between the facts of passing the module and the contents of each one.

In view of the above, it indicates that the content of the modules was not the variable which determined the passing scores. This was true whether the passing score was MB, B, or S.

**DISCUSSION**

Meaningful learning theory is based on the construction of the knowledge by the students (Ausubel, 2000; Bretz, 2001), with this idea in mind; three strategies were implemented by the professor in each module:

1. The learning-teaching process was performed with a gender equity perspective, with the same level of participation for both the genders.
2. Molecular models were used by the students in class to achieve the spatial skill necessary to understand the stereochemistry topic, which is of the utmost importance to distinguish and discern the reactivity and...
transformation of the molecules in module O. For module P, they had practiced drawing flow diagrams.

3. Applying open-book examinations with open questions.

Each theoretical session included addressing and discussing the items of the homework by the students, teaching of new concepts by the professor, and topics or reactions teaching by the students with examples selected by them. The professor encouraged the students to use other exercises than those from the books, and the students should actively participate at least twice in each theoretical session.

The professor, herself, had worked with molecular models during her master degree in the course of organic chemistry and she felt that this experience was very beneficial. Her own positive experiences were the motivation to use molecular models. These were to help her bachelor students develop their own spatial skills. She observed that her students were more engaged when they handle the molecular models. Previous research has reported that this strategy has helped to develop the spatial ability of both female and male students (Bodner et al., 2003; Sorby, 2009).

The third strategy of open-book examinations was used by the professor because it was also a satisfactory experience during her bachelor degree. The open-book examinations caused anxiety among the students because they thought that this type of test would be more difficult than the close-book examinations. In addition, they also reported that they were accustomed to memorizing the information expected on close-book examinations. At the end of each module, they noted that the difficulty of the examinations was comparable.

In relation to the types of examinations, Sato et al. (2015) found that in biology there was no significant difference between the groups of students who took open-book and closed-book tests. The same result was reported by Brightwell et al. (2004) for anatomy. Heijne-Penninga (2010) highlighted that open-book exams were not suitable for medical students at Groningen University in The Netherlands. However, to the best of our knowledge, there is one paper that has dealt with the differences between multiple-choice and open response tests in organic chemistry (Tang and Xie, 2017) but not any concerning open-book examination in organic chemistry or physical chemistry as is presented in this work. While this was true for biology (Sato et al., 2015) and anatomy (Brightwell et al., 2004), Heijne-Penninga (2010) highlighted that open-book examinations were not suitable for medical students at Groningen University in Netherlands. However, to the best of our knowledge, there are papers that have dealt with the differences between multiple choice test and open response in organic chemistry (Tang and Xie, 2017) but not any concerning open-book examinations in organic chemistry or physical chemistry as is presented in this work.

It was not possible to compare our results with previous studies because of the types and number of variables that has been taken in account on all of them; for example, Shibley et al. (2003) found a lack of a significant gender difference in the total inventory Piaget’s developmental task score applied in an introductory chemistry course with students who were not in science programs. Meanwhile, Turner and Lindsay (2003) had analyzed scores of two groups of an introductory course of organic chemistry, where they had searched whether gender differences existed for cognitive skills and emotional responses. They had suggested that substantial gender differences in cognitive and non-cognitive factors related to achievement in organic chemistry may be found at the level of college science courses. However, those groups had two different instructors with different type of tests.

In other articles, even though the research dealt with organic chemistry, the effect of genus was not considered. Trying to understand, how and why a group of undergraduate students chose either meaningful learning or rote memorization in organic chemistry, Grove and Bretz (2012) interviewed participants and analysed their written essays. Meanwhile, Dori and Barak (2000) had found by means of questionnaire that students from high schools in Israel that were taught with three-dimensional molecular models understood the model concept better and were capable of applying transformation from one-dimensional to two- or three-dimensional molecular representations than students who did not have that experience in class. However, there was no breakdown, to itemize by gender.

The above-mentioned studies have revealed the size and complexity of the subject, but it is worth the effort to find the better conditions which allow to improve the performance of the students in organic chemistry and physical chemistry for both genders. Our data showed that in module O, where the spatial skill is very important together with reasoning, females had some better outcome than males, which was in contrast with the gender differences in favor of men related to the spatial skill reported by Nordvik and Amponsah (1998).

In the present case study, the teaching strategies used in both modules, as stated, reflect meaningful learning theory as they emphasised reasoning and the construction of knowledge. Our results seemed to fit, to some extension with Ayalon and Livneh (2013), who had found that countries that expose their students to similar math knowledge were characterized by lower levels of gender inequality, even though they had analyzed population of fourth and eighth graders based on the trends of the International Mathematics and Science Study 2003 for different countries.

We know that there are a lot of studies about the achievement in mathematics, and for instance, Ross et al. (2012) had given a list of them in which the gender differences exist and those which reported that gender differences in mathematics achievement are near zero as Ross found for Canadian students of secondary school. The question is, therefore, which were the circumstances and variables that promoted the equality in sciences? The results obtained in module P, where the use of
physics and mathematics are necessary, back the statement that there was no association between the scores and gender using our strategy of teaching.

**CONCLUSION**

The content of the module was not the variable which determined the passing score in any of the categories MB, B, and S. There were other variables that could be considered to increase terminal efficiency such as harmonization of the background knowledge, spacious classrooms, or improvement of comprehending texts skill. There were no statistical differences between the proportions of failing scores per module. The failing rate and the low terminal efficiency could be due to that the students were not accustomed to answering open-book examinations. Some of them decided to look for the answers in their books or module notes, whereas others seemed not to be able to justify their answers or sketch the flow diagram. We think that the open-book test must be promoted in chemistry because it can allow the students to become decision makers, and real life requires a more open mind and skills to analyze information rather than just memorizing data.

Association between the scores and gender was found for the O subgroup (p = 0.04) but not for P subgroup (p = 0.26). These results pointed out that undergraduate female students had a better performance in organic chemistry than the male students and that both genders showed the same success on the physical chemistry subjects of the module P. This result suggests that the gender equity participation in the sessions, the use of molecular models to promote spatial skills, and open-book examinations could help the female students to learn chemistry and physical-chemistry with better or similar outcomes to their male counterparts.

**ACKNOWLEDGMENTS**

The authors would like to thank Victoria Alcántara Gutiérrez for the technical assistance.

**REFERENCES**


