Results analysis between embedded assessments in Semestre i versus isolated assessments in education.

Reséndez-Maqueda, Luis Fernando[1]; Sandoval-Correa, Alejandro[2]; Forte-Celaya, Maria Rubi[3]
(1): luis.resendez@tec.mx; (2): alesando@tec.mx; (3): mrforte@tec.mx
Escuela de Ingeniería y Ciencias, Tecnológico de Monterrey, Bulevar Enrique Mazón 965, Sonora, México.

Abstract

The educational model of the Tecnológico de Monterrey (México) is on transition; it focuses on a new way of organizing and evaluating the teaching-learning process. The TEC21 educative model is an academic exercise that strives on the development of competences based on solving challenges (ITESM, 2015). Our project focuses on presenting the preliminary results of an integrated evaluation of a semester (*Semestre i*) versus a final integrated evaluation. During the last semester of industrial and systems engineering undergraduate program, students at Tecnológico de Monterrey participated in a Final Assessment in order to measure the development of competences and skills corresponding to their profession. They solved a problem under time pressure conditions, external irruptions, theoretical and technical requirements that will be present in their future work situations.

On the other hand, the exercise called “Semestre i” focuses on the student's relationship with the work environment, through solving challenges, oriented to develop competences (ITESM, 2017). Students live a continuous embedded assessment within the resolution of a designated challenge through 18 weeks, monitored and evaluated by different stakeholders. It is important to highlight that results are presented periodically during the semester by the students which are under constant surveillance, working in teams and managing time and resources.

In the final assessment, the results tend to be higher and with lower standard deviation when grading competences development. While the evaluation included in “Semester i” shows an increment in the differences between the student outcomes and with greater standard deviation. Integrated assessment has a natural "tuning" of the criteria that must be considered, with a particular emphasis on technical aspects. Such considerations tend to be ignored in the isolated assessment, where soft skills seem to be the most valuable competences considered by the evaluators.

Keywords: Assessment, Competencies, Evaluation, Continuous Improvement

Introduction

Tecnológico de Moterrey's mission is *"We develop leaders with an entrepreneurial spirit, a humanistic outlook and a global vision"* ITESM SP (2015) . This mission is completed through student's competencies development. Competencies are defined as the conscious integration of knowledge, abilities, attitudes and values, ITESM (2018). In order to validate such development, assessment exercises has being deployed along undergraduate students curriculum. However different assessment approaches show inconsistent results regarding students historical performance. In some of the assessment exercises, considered good students (according to his/her historical performance) may show results that places them in a lower level of development of competencies, meanwhile students considered inconstant (according to his/her historical performance) may show results that places them in a higher level of development of competencies. Such results not only give feedback to students but also are an important source of information for curriculum continuous improvement process. It is then essential to deploy assessment exercises, whose characteristics leads to accurate results.
In the next sections we explore theoretical approaches that reinforces our proposal, based on the experiences acquired performing different assessment exercises in the frame of Tec21 educative model.

Literature Framework

According to Jacobs et.al (2018), the assessment center permits the competency level evaluation of the participants, which need to solve problems in a similar business environment where observers can see their performance and identify strengths and opportunities for improvements.
The main objective of assessment center is to evaluate the participants in the key characteristics and competencies that enterprises want to observe in a potential collaborator. Commonly this type of exercises are held in a few hours, in a simulated environment and with the conscious of being observed. All those characteristics can interfere with the participants natural behavior, who may influence results. People trends to give the best impression and try to do what they think is expected by the evaluators/observers, influencing their opinion. This behavior at first was described by Erving Goffman as impression management in 1959, and Klehe et.al (2014) mentioned that this behavior is commonly seen and increased by participant in the assessment center exercises rather than in other no evaluated and observable situations, because people goes with the main objective of being noticed and pursued a great evaluation and the desirable outcome.

The observer effect was very well documented for Elton Mayo, throughout the famous Hawthorne effect, talking about performance in a production plant. However this effect is not only exclusive for production plants, Jonathan S Myers et al (2016) presents a study about Impossibility to eliminate observer effect in the assessment of adherence in clinical trials. Myers et al (2016) Shows that an observer effect was not reduced,in three different variants in subjects adhered to taking their medications. As in medical treatments, in assessments conceived and announced to evaluate individual participating this observer effect produces an artificial behaviour that will disappear once the participants will not be longer monitored. Talking about assessment of competencies development, it is expected that individuals will behave in the observed way in any other similar circumstances either he or she will be observed or not.

According to Randy E. Bennett, one important aspect that should include educational assessments in the future should be the Greater Use of More Complex Tasks, he explains that the activities that characterize proficiency in a discipline often take the form of extended problem-solving episodes. Making a comparison with the new Educational Model proposed by Tecnológico de Monterrey, we can find that the students are located at the center of the model and that through solving complex problems or challenges, they will acquire knowledge and develop competencies.

Specifically talking about the *Semestre i* program, the way of assessing the knowledge acquisition and competence development is through the constant evaluation of academic content and by solving real life problems inside an organization. Students need to use all the technical knowledge and tools learned in the classroom, and apply them in the resolution of a real-life case. The length of the Semestre i program is 18 weeks, so this time frame allows to have enough time for the students to analyze and understand the problem to solve, to define the most accurate solutions to it and to start an early stage of implementation of these solutions.This process would be almost impossible to repeat in a classic assessment session, where the time is limited and the tasks to be performed usually are not very complex.

According to Foss (2003), the tacit and socially embedded aspects of knowledge may come over the limits established by the bounded rationality. In our assessment design, the level of complexity of the situation that students overlook may force them to face challenges if a number of competencies has been transformed into tacit knowledge. Therefore complexity must be present in an assessment exercise, in order to force students to respond in a more natural and spontaneous way if the competencies are already embedded in their action-reaction reflex.

Evaluators in a *Semestre i Embedded assessment* track students the long of a semester as students develop the project (challenge). Evaluators in the final Programme Assessment discover students the day of the evaluation. Maria de Lurdes Calisto (2017) describe some behaviours talking about Senior-level Managers assessing employees talking about intrapreneurial behaviour. *Calisto* suggests that managers base their decisions on a fragment or subset of available information. According to Calisto, individual bounded rationality might be compensated by the short social distance. This argument is based on the assumption that contact between socially proximate individuals occurs at a higher rate than among socially distant individuals. Following with Calisto's analysis, is the idea that social imitation process may help people make decisions with limited knowledge. It might be that senior-level managers base their assessment of employees on the clues they pick from other people within the firm, for instance middle-level managers. And finally, Calisto ads a third explanation might have to do with the nature of corporate entrepreneurship (CE) itself. CE might be more clearly understand under a social exchange perspective, more than under the bounded rationality perspective. A social exchange perspective highlights CE's on-going, dynamic quality. Individual actions and decisions are thus seen in a relational context, suggesting that managers are continuously interacting with, and learning about, employees. This third Calisto's argument matches in a better way with the assessments that happens the long of a *Semestre i*, in contrast with what happened in a Final Programme Assessment, where we can identify more with Calisto's point 1 and 2.

Dietel R.J et. al (2003) stated in his article "What Does Research Say About Assessment?" that an important characteristic of a good assessment is that it can be reliable and offer validity depending on the consistency of the results obtained. In the case of the Embedded Assessment results are

consistent because the evaluation is performed in a periodic manner and it did not show significant variations. On the other hand, Final Programme Assessment is an isolated event and also it has a limited time frame to perform the evaluation..

In the same order of ideas, Mohammad AL- Shehri et al. (2015) refers that the authentic assessment strategies has the best results by reinforcing actual learning by connecting students' learning with real life challenges and triggering students critical thinking. Al Shehri et al, points in how useful are actually assessment in which a student can be involved in certain activities to get more information about student's ability to apply what he has learned by using new and various situations.

Analysis

Considering the elements detailed in the Theoretical Framework and linking them with the assessment initiatives performed at Tecnológico de Monterrey, the proposal of this article can be focused on comparing the *Semestre i* exercises versus the final programme assessments executed during the last academic periods. The main premise is that the grades obtained by the students during the assessment are in general higher than the ones obtained during a complete semester of being evaluated (embedded assessment) observed in a *Semestre i*.

First, we will start with describing the methodology followed on each exercise in order to have a general context and then we will present the grades obtained by the students on each one of them.

Semestre i Projects

The *Semestre i* is an academic program developed by Tecnológico de Monterrey and it is part of the new Educational Model Tec21. This format consists of developing an academic semester that combines experiential learning, theoretical modules, and challenges faced inside a company that participates as a formative partner.

The students fulfill the set of subjects that they have to study during their 5th semester, and also develop disciplinary competencies facing problems and challenges in a real environment. Students work intensively during 18 weeks, the first 6 weeks of the semester inside the classrooms, and during the following 12 weeks they experience a period of total immersion inside the company.

In March 2016, an agreement was signed between the company BOSCH Security Systems, located in Hermosillo, Sonora, and the Tecnológico de Monterrey Campus Sonora Norte, to jointly execute a *Semestre i* dedicated to both academic and competence development for students of the fifth semester of the program Industrial and Systems Engineering (IIS).

Four teams of students were formed, they were supported by an ITESM Tutor, and a BOSCH engineer who served as a Mentor. The role of the Tutor was to monitor student learning and validate the development of Industrial Engineering disciplinary competences. The Mentor role was assigned to the Bosch members who were responsible of the areas where the projects were being executed, and their main responsibility was to be the linkage between the student and the operators of the area, they were also in charge of assisting the students and help them to become familiar

with the Bosch working method, the formats, reports, and technical language used in the plant, as well as to guide them in the planning and daily execution of the tasks.

Additionally, the intervention of the Plant Manager was very important and decisive for the correct execution of the projects. During the weekly meetings that the management team had in Bosch, they assigned 15 minutes to the student teams for presenting their progress and receive direct feedback about their contributions. Undoubtedly, in these sessions the students faced their first big challenge since they had to be very precise and concise when presenting the information, also they needed to be very receptive to constructive criticism and the evaluation of their performance.

The structure of the Semestre i can be appreciated graphically in the following figure, where the coloured boxes represent theoretical modules and the yellow sections the time assigned to solve the challenges:

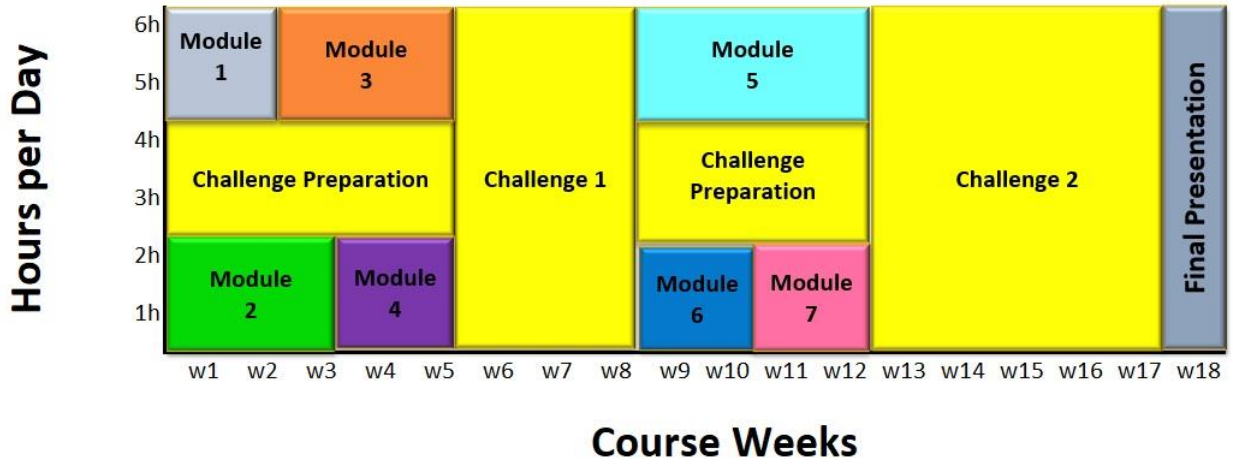

Figure 1. Distribution of academic modules, preparation and challenge times, of *Semestre i* "Industrial Engineering Squad in BOSCH".

The evaluation instruments used during the *Semestre i* were not only exams or practices, but also there were used individual logs, essays, technical reports, and presentations. We can then talk about an Embedded Assessment. The application of these instruments required a close accompaniment of the teachers, who conducted weekly interviews with each team member, to ensure progress in the development of skills, and give the proper feedback to the students.

Also, a percentage of the evaluation was provided by the mentors of the company, they were in constant communication with the students and they were validating the solutions and improvements during all the project.

Final Programme Assessment

The Final Programme Assessment consists of a 4 hour session where the senior students have to solve an academic case related to Industrial Engineering and their performance is observed by experts and employers from the most important companies of the region. The session is divided into three main moments, the first one includes an introduction to the case for the students and

their initial analysis of it. The second moment allows the students to decide which concepts or Industrial Engineering tools should be used to solve the case in the most efficient and optimal way. And finally, the third moment where the results are presented to the evaluators and they provide individual feedback to the students about their performance and recommendations of personal improvement.

For this exercise, the evaluation was made by the experts and employers that were previously defined, they did not have any prior information about the students in order to avoid any kind of prejudice or misconception.

In Table 1, we map the characteristics observed in each type of Assessment, based on our literature framework. In such table, beyond the duration, we pound the level of each Characteristic's influence, as Weak, Medium or Strong.

Characteristics	**Embedded Assessment**	**Final Programme Assessment**
Length	18 weeks	4 hours
Observer effect	Medium	Strong
Reliability and validity	Strong	Medium
Social distance	Weak	Strong
Social imitation	Weak	Strong
Social exchange	Strong	Weak

Table 1. Characteristics of Embedded Assessment and Final Programme Assessment

Results Analysis

Using the evaluation results obtained from both the final presentation in the Embedded assessment (*Semestre i*), and the Final Programme Assessment's, we present a statistical analysis, comparing both exercises.

In the case of the embedded assessment we are using a set of 15 students' results. For the final Programme Assessment, we are counting a total of 24 registers. In both cases we are considering five aspects to evaluate: 1) Problem identification, 2) Diagnosis, 3) Action plan, 4) Knowledge, and 5) Presentation.
It is interesting to remark that 13 students were involved in both assessment exercises.

Data sample was analysed with the statistical software Minitab, first of all, descriptive statistics were obtained for each aspect in both exercises, Final Programme Assessment results are Type A and embedded assessment (*Semestre i*) results are Type I, as shown in Figure 2.

Descriptive Statistics: Ident, Diag, Act, Know, Present, Average

Variable	Type	Total Count	Mean	StDev	Variance
Ident	A	24	2.7083	0.4149	0.1721
	I	15	2.600	0.632	0.400
Diag	A	24	2.6042	0.4099	0.1680
	I	15	2.400	0.632	0.400
Act	A	24	2.5625	0.4499	0.2024
	I	15	2.667	0.488	0.238
Know	A	24	2.8542	0.2750	0.0756
	I	15	2.533	0.640	0.410
Present	A	24	2.8125	0.3234	0.1046
	I	15	2.467	0.516	0.267
Average	A	24	2.7083	0.2753	0.0758
	I	15	2.533	0.445	0.198

Figure 2. Descriptive Statistics for Type A and I results

According to information obtained for this *first case*, in all aspects, bigger variances identified, are from Semestre i and the bigger means are in Final Programme Assessment exercise. Also it can be inferred that means are similar, but variances might be different in some aspects. As shown in the Boxplot in Figure 3, when comparing Type A and Type I assessments, the dispersion is particularly different for the *Identification*, *knowledge* and *Presentation*, also the levels' average seems to be different as well.

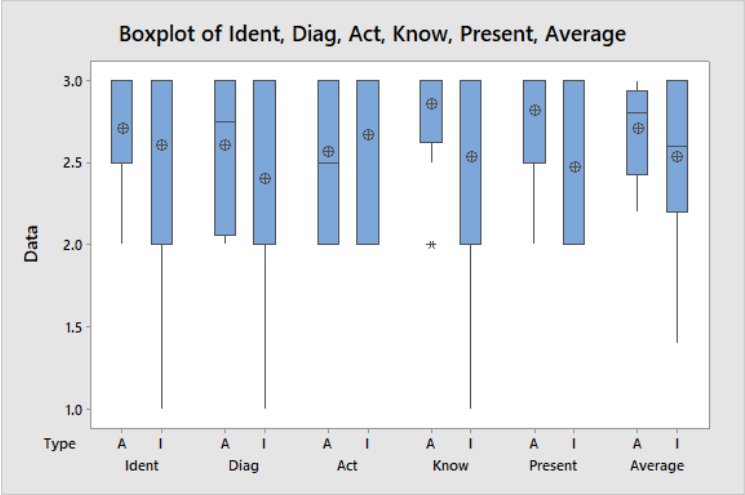

Figure 3. Boxplot for Results Assessment Type A and Type I

Levene's method was used to compare variances of both exercises, type A and type I, due to the sample size and distribution.

Initially, our first hypothesis was that the average variance would be significantly different when comparing both exercises, however the two variances test showed that the average variance was not considered statistically different (Figure 4).

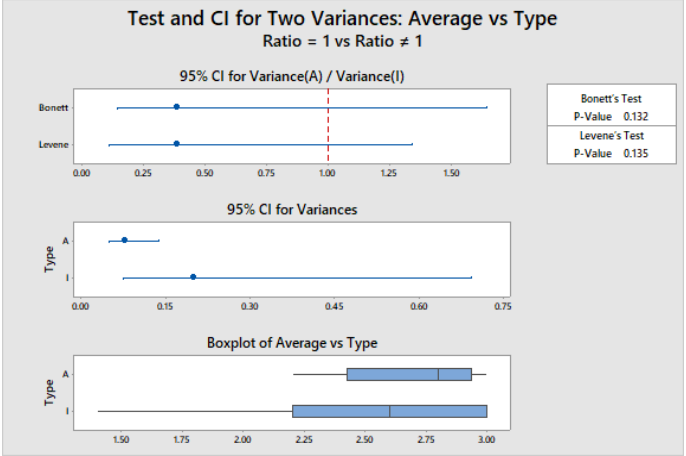

Figure 4. Test for two variances: Average for Type A and Type I

As a next step, we deployed an analysis splitting the five different aspects. The variance's analysis shows that only *Knowledge* (Figure 5) and *Presentation* (Figure 6) are significantly different.

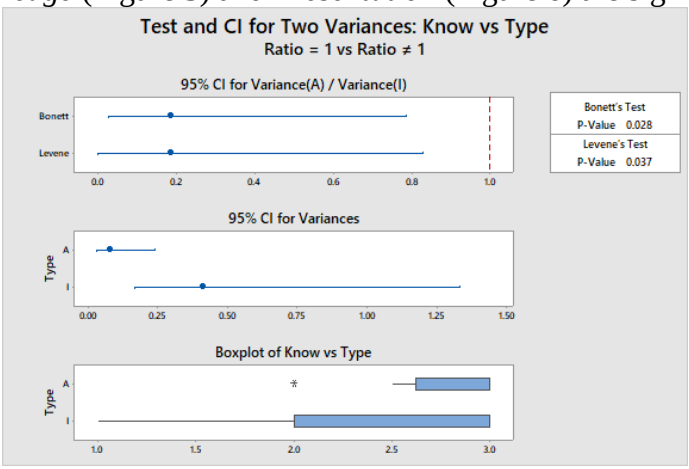

Figure 5. Test for two variances: *Knowledge* for Type A and Type I

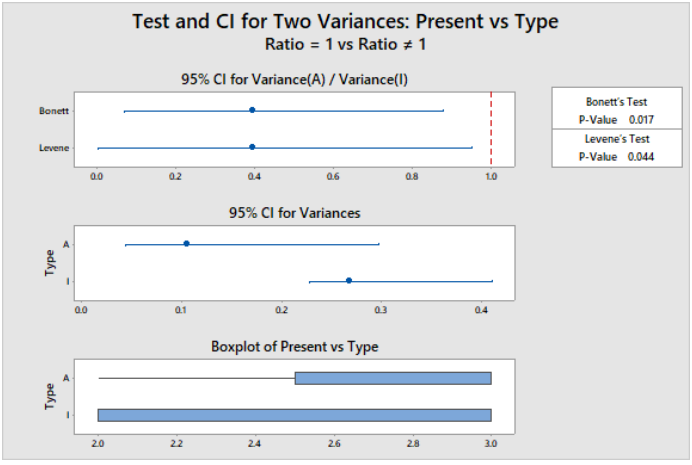

Figure 6. Test for two variances: *Presentation* for Type A and Type I

Conclusion and discussion

Regarding Figure 4, the *Average* comparison of both exercises, even when the difference is not yet significant, it is still interesting to observe the level of dispersion. A new analysis containing a greater number of elements in a sample, as well as a more accurate election of aspects to evaluate, may reveal new precision in the future.

Regarding Figure 5, talking about *Knowledge,* it is confirmed that a significant difference between variance in results from both exercises was found. Additionally, an important dispersion range is observed in the Embedded Assessment, while in the Final Programme Assessment the range suggested a greater homogeneity in the competency evaluation level.

Regarding Figure 6, talking about *Presentation,* it is also confirmed that a significant difference between variance in results from both exercises was found. In this case, as well, an important dispersion range is observed in the Embedded Assessment, while in the Final Programme Assessment the range suggested a greater homogeneity in the competency evaluation level.

According to these results, it is suggested that the Embedded Assessment's features allow more differentiation within the performance levels of the participants that are being evaluated.

Since the participants of this evaluation process are particularly undergraduate students, the main objective is to provide feedback about the present status of the competencies development and gives academics valuable information in order to achieve continuous improvement. So, according to the present analysis, Embedded Assessment seems to be the more suitable exercise.

References

Bennett, Randy E. Educational Assessment: What to Watch in a Rapidly Changing World. Educational Measurement: Issues and Practice Winter 2018, Vol. 37, No. 4, pp. 7–15

Dietel, R.J., Herman, J.L., & Knuth, R. (2003). What Does Research Say About Assessment.

Eraut, M. (1994). Developing Professional Knowledge And Competence. London: Routledge. Retrieved from http://0-search.ebscohost.com.millenium.itesm.mx/login.aspx?direct=true&db=nlebk&AN=82987&site=eds-live

ITESM (2015) Modelo de Programas Formativos de Profesional TEC21. D.R.©, Instituto Tecnológico y de Estudios Superiores de Monterrey, Av. Eugenio Garza Sada Sur No. 2501, C.P. 64849, Monterrey, N.L., 2015.

ITESM SP (2015) The Tecnológico de Monterrey's 2020 Strategic Plan©. Instituto Tecnológico y de Estudios Superiores de Monterrey, Av. Eugenio Garza Sada Sur No. 2501, Monterrey, NL, Mexico, 64849, 2015.

ITESM (2018) Modelo educativo Tec21 (Tec21 Educative model). Instituto Tecnológico y de Estudios Superiores de Monterrey, Av. Eugenio Garza Sada Sur No. 2501, Monterrey, NL, Mexico, 64849, 2018.

Jacobs, R. R., Griswold, K. R., Swigart, K. L., Loviscky, G. E., & Heinen, R. L. (2018). From Campus to Corporation: Using Developmental Assessment Centers to Facilitate Students' Next Career Steps. *Journal of the National Collegiate Honors Council*, *19*(1), 125–154. Retrieved from http://0-search.ebscohost.com.millenium.itesm.mx/login.aspx?direct=true&db=eric&AN=EJ1185813&lang=es&site=eds-live

Jonathan S Myers et al (2016). Impossibility to eliminate observer effect in the assessment of adherence in clinical trials. Patient Preference & Adherence. Oct2016, Vol. 10, p2145-2150. 6p.

Klehe, U.-C., Kleinmann, M., Nieß, C., & Grazi, J. (2014). Impression Management Behavior in Assessment Centers: Artificial Behavior or Much Ado About Nothing? *Human Performance*, *27*(1), 1–24. https://0-doi-org.millenium.itesm.mx/10.1080/08959285.2013.854365

Maria de Lurdes Calisto (2017). Senior-level Managers' Traits and the Assessment of Intrapreneurial Behaviour – Exploring the Relevance of Bounded Rationality. Estoril Higher Institute for Tourism and Hotel Studies, Portugal Center for Advanced Studies in Management and Economics of the Universidade de Évora (CEFAGE-UE)

Mohammad AL- Shehri et al. (2015), The Range of the Availability and Practice of Authentic Assessment Competencies for Intermediate Stage Teachers in Sharurah Governorate. Journal of Education and Practice www.iiste.org ISSN 2222-1735 (Paper) ISSN 2222-288X (Online) Vol.6, No.32, 2015

Nicolai J. Foss (2003). Bounded rationality and tacit knowledge in the organizational capabilities approach: an assessment and a re-evaluation. Industrial and Corporate Change, Volume 12, Number 2, pp. 185–201