

Evaluative revolution in university education in the age of social networking

Edwin A. Hernández¹, Humberto J. Navarro², Omar Lengerke², Andrés M. García¹

¹ Faculty of Socioeconomic and Business Sciences, Unidades Tecnológicas de Santander, Bucaramanga, Colombia

² Faculty of Natural Sciences and Engineering, Unidades Tecnológicas de Santander, Bucaramanga, Colombia

*Corresponding author E-mail: ehernandez@correo.uts.edu.co

ABSTRACT

Evaluation mechanisms are a subject of analysis to strengthen the performance of university students; this study compares two modalities: evaluation through activities on the social network Facebook and assessment through a multiple-choice test. Using a quantitative methodological design with a comparative approach between the two models, grades were collected and entered into the SPSS 26 program as a first step. The study used mean analysis tools such as ANOVA, Levene's test, and Post Hoc (Games-Howell) to measure significant differences between means and variances of the grades of three academic cuts. The target sample consisted of two groups of 44 students from the Unidades Tecnológicas de Santander. The results show that the Facebook evaluation model did not present significant differences between the cuts, which is inferred as homogeneity in the grades. On the other hand, the written evaluation process showed substantial differences between the grades of each cut. As a main finding, it is concluded that the type of evaluation impacts the grades. The dialogic measurement by social networks allows for equity in the assessment of the group, while the written tests generate a differentiation in the students' performance.

Keywords: student assessment, social networks, multiple response test, higher education, comparative analysis

1. Introduction

The knowledge evaluation process has been considerably reformulated over the years, mainly focusing on the competencies acquired by university students in the course development, and with the rise of Information and Communication Technologies (ICT) and Web 2.0. In the educational context, the evaluation model has been criticized for its limited approach that privileges memorization over skills such as innovation, problem solving, leadership, and teamwork [1]. On the contrary, recent models promote an assessment approach linked to the reality of the needs of students who are about to enter the labor market, among the main deliverables are problems that seek a solution in teams, the analysis of a business case or the presentation of a research project [2], [3]. Social networks play an essential role in the transformation of learning environments, becoming a fast and effective two-way knowledge transfer mechanism, which is manifested by the continuous and active interaction of students and teachers using mechanisms and applications, engagement, closed group management, and chat publications [4], [5], [6], [7], [8]. In the case of the Facebook social network, the evaluation process of the knowledge and skills acquired by students is guaranteed by the dialogic feedback of the various deliverables by classmates, teacher, and self-evaluation. In this sense, integrating social networks as a mechanism for evaluating knowledge and content in higher education students is based on theories such as constructivism, which holds that learning is based on mediation, feedback, and interaction with the social and cultural environment [9]. From this point of view, the social network Facebook contributes as a facilitator of collaborative learning and self-learning. On the other hand, the written test continues to be developed in the classroom, even though it has been accused of superficiality in learning, in the same way, it can be detached from the construction of a product or deliverable improved and perfected by the student and generate high levels of stress and anxiety [10]. Nevertheless, it has advantages, such as the ease with which the teacher can grade it and the measurement of conceptual knowledge [11].

From this perspective, García-Ruiz et al. [12] state that the learning environment influences grades and the developed skills. In that order of ideas, the need arises to analyze various assessment modalities, which integrate Web 2.0, specifically social networks, to impact performance and parity in academic scoring [13], [14]. Despite



this, the change of traditional assessment practices generates a series of challenges that involve resistance to change by teachers, lack of rigor in the correction, and distractions that the use of Facebook [15], [16], [17]. Therefore, research that equates two study variables is necessary to understand how different learning environments affect knowledge, competencies, and grades [18].

This study aims to analyze how the type of evaluation and the learning environment can significantly influence academic performance and the dispersion of students' grades [19], [20]. In this way, it is proposed to carry out, on the one hand, a process that consists of assigning a deliverable, with a series of requirements in a sample of 44 fifth semester students of the academic program Technology in Business Management of the Technological Units of Santander; the deliverable is subjected to dialogic evaluation firstly, by classmates and secondly, by the teacher, in both cases they are evaluated in the social network Facebook. To compare the impact of an evaluation based on creating a deliverable on Facebook. On the other hand, the assessment uses a questionnaire with multiple-choice questions applied to another sample of students of the same academic semester; this evaluation assesses the knowledge and skills learned in the scholastic period. The study allowed for evidence that supports the idea that social networks, specifically Facebook, can be an effective tool to facilitate dialogic feedback and maintain a stable level of performance in the quality of the deliverables presented by the students. Additionally, the written test could have contributed to increasing the dispersion in performance, promoting less consistency in the results.

2. Research method

2.1. Type of research

The study is descriptive with a quantitative approach, analyzing grades divided into three academic cuts. Specifically, a descriptive analysis, Levene's test, and ANOVA were performed for the study with deliverables on the social network Facebook because the variance of the three cuts is similar. In contrast, descriptive analysis, Levene's test, and Post Hoc (Games-Howell) are performed for the written test due to the significant differences between variances generated.

2.2. Sample

The research sample consisted of 88 Business Management Technology Program students at a university in Santander, Colombia. The research participants were divided into two groups, grouped as follows: on the one hand, 44 students carried out the evaluation exercise through the social network Facebook, and applied the dialogic feedback with their peers or classmates. This group of students was evaluated with deliverables presented and assessed through the selected social network. On the other hand, 44 students followed the traditional model of written and face-to-face evaluation with a grade established by the teacher; this study method allowed contrasting the performance of students evaluated by Facebook and those evaluated face-to-face. Quantitative data from assessments in three academic periods were analyzed to identify trends in student performance and their possible relationship with the use of social networks.

2.3. Data analysis

An assessment of the grades of three cuts was carried out on a sample of 88 university students. Once all the grades were collected, they were entered into the SPSS 26 program to perform descriptive analyses such as frequency, standard deviation, and ANOVA to measure whether there were significant differences in the grades of the three cuts. Likewise, several null and alternative hypotheses analyses were performed, evaluating the F-statistic and the p-value.

3. Results and discussion

3.1. Evaluation by the social network Facebook

The results of the grades of the three cut students who carried out their formative process through the social network Facebook and who were evaluated using a dialogic evaluation of a deliverable presented by the social network under study are presented.

3.1.1. Descriptive analysis

Firstly, Table 1 presents descriptive data of the study, including the mean of the three cuts, the standard deviation, the confidence interval, and the comparison graphs between the three means.

Table 1. Descriptive results of the Facebook evaluation model

Rating	N	Media	Deviation	Error	95% CI		Minimum	Maximum
					Lower limit	Upper limit		
1	44	4,316	0,4993	0,0753	4,164	4,468	3,2	5,0
2	44	4,368	0,5548	0,0836	4,200	4,537	3,0	5,0
3	44	4,289	0,4065	0,0613	4,165	4,412	3,2	5,0
Total	132	4,324	0,4880	0,0425	4,240	4,408	3,0	5,0

Note: 95% CI = 95% confidence interval for the mean.

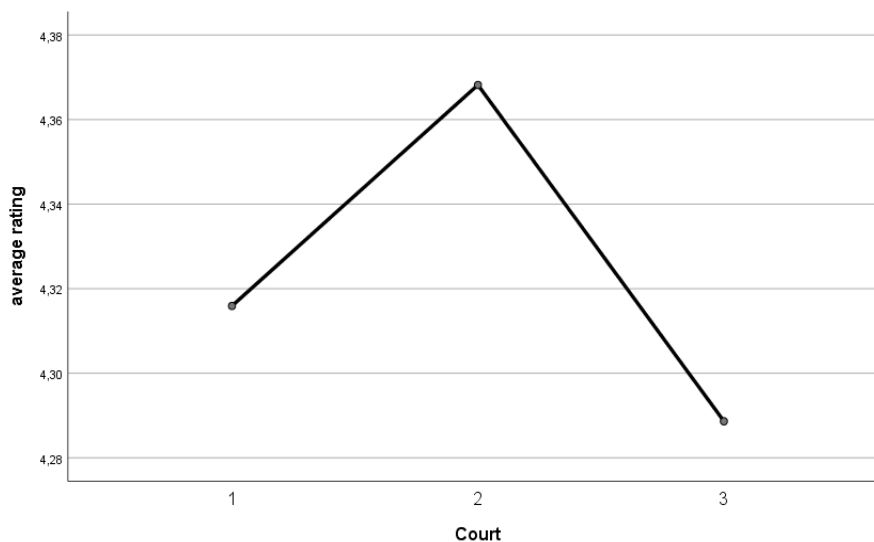


Figure 1. Graph of the average score of the three academic cuts of the Facebook evaluation model

According to the mean analysis, the second cut had the highest mean (4.368), and the third cut was the lowest (4.289); however, the differences between them are insignificant. The standard deviation shows greater variability in the second cut-off (0.5548), which implies a greater difference in the grades in this cut-off. Finally, the confidence interval (95%) is analyzed, and all confidence intervals overlap, which supports that there are no statistically significant differences between the means, thus confirming the need for ANOVA. The null and alternative hypotheses are presented before the homogeneity of variances test is presented.

Hypothesis:

- H₀: The variances of the ratings among the three groups (cut-offs) are equal.
- H₁: At least one of the variances is different.

Table 2. Levene's test, Homogeneity of variances test

		Levene's statistic	gl1	gl2	Sig.
Rating	Based on the average	2,314	2	129	0,103
	Based on the median	2,063	2	129	0,131
	It is based on the median and with adjusted gl	2,063	2	122,881	0,131
	It is based on the trimmed mean	2,100	2	129	0,127

In evaluating the null hypothesis (H₀) that the variances between the groups are equal, the p-value (sig) is greater than 0.05 in all cases. This allows us to determine that the null hypothesis is not rejected and that the assumption of homogeneity of variances is fulfilled, which validates the use of ANOVA.

Table 3. Anova

Rating	Sum of squares	gl	Root mean square	F	Sig.
Between groups	0,144	2	0,072	0,299	0,742
Within groups	31,059	129	0,241		
Total	31,202	131			

The ANOVA table shows that the F value = 0.299 indicates a minimal variation between group means compared to the within-group variation. Meanwhile, the p-value = 0.742 is much greater than 0.05, so t. Therefore, the thesis is not rejected, and there is no statistically significant difference between the grade means of the three cuts. Finally, it is essential to note that since there are no significant differences between the means, the post hoc test is not performed.

The results infer that formative processes mediated and evaluated by the social network Facebook allow for stability in the students' grades; statistically, the analyses between the means were carried out, and there are no significant differences in the final grades of the three academic cuts.

3.2. Results of evaluation using written and face-to-face evaluations

The grades of the three cut students who completed their formative process by submitting a written evaluation face-to-face following the traditional classroom method are presented.

3.2.1. Descriptive analysis

Firstly, the descriptive data of the study are presented, in which the mean of the three cuts, the standard deviation, the confidence interval, and the comparison graphs between the three means are analyzed.

Table 4. Descriptive results of the written evaluation model

Rating	N	Media	Deviation	Error	95% CI		Minimum	Maximum
					Lower limit	Upper limit		
1	44	2,573	1,0750	0,1621	2,246	2,900	0,6	5,0
2	44	3,739	1,0814	0,1630	3,410	4,067	1,0	5,0
3	44	4,211	0,2687	0,0405	4,130	4,293	3,3	4,5
Total	132	3,508	1,1246	0,0979	3,314	3,701	0,6	5,0

Note: 95% CI = 95% confidence interval for the mean.

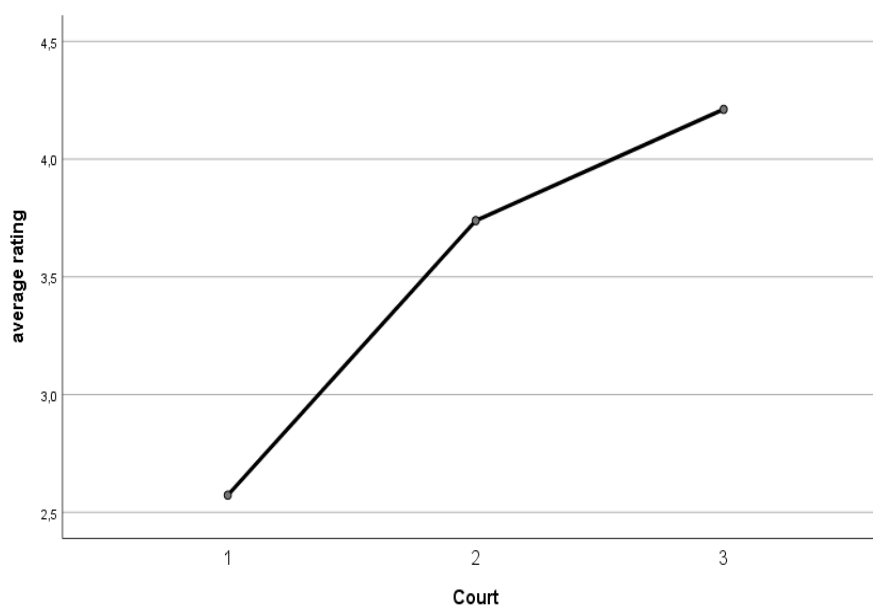


Figure 2. Graph the average grade of the three academic cuts of the written test evaluation model

The analysis of the averages shows that the first cut had the lowest average (2.573) and a progressive improvement in the average grade, reaching the third cut with the highest average (4.211). Likewise, a clear

difference can be observed between the three cuts, specifically between cut one and the other two, as detailed in the graph of the average of the three cuts.

The standard deviation shows greater dispersion in cuts 1 and 2, with values above one point, and a third cut with a slight variability of (0.2687), which allows inferring that those students obtained similar grades.

Finally, we analyzed the confidence interval (95%) and found that the intervals do not overlap. This supports the statistically significant differences between the means, which allows us to interpret that an ANOVA would confirm substantial differences.

The null and alternative hypotheses are presented before the homogeneity of variances test is presented.

Hypothesis

- H_0 : The variances of the ratings among the three groups (cut-offs) are equal.
- H_1 : At least one of the variances is different.

Table 5. Levene's test, Homogeneity of variances test

		Levene's statistic	gl1	gl2	Sig.
Rating	Based on the average	22,287	2	129	0,000
	Based on the median	15,743	2	129	0,000
	It is based on the median and with adjusted gl	15,743	2	90,641	0,000
	It is based on the trimmed mean	20,911	2	129	0,000

All methods yield a result $p = 0.000$, less than 0.05. This means the null hypothesis of equality of variances is rejected; therefore, the variances between the groups are significantly different.

The results allow inferring that formative processes, evaluated with a written test, do not stabilize the students' grades. Statistically, the analysis between the means was carried out, and significant differences are presented in the final grades of the three academic cuts, with a minimum of 0.6 and a maximum of 5 in the first cut.

To deepen the statistical analysis, specifically the behavior in each cut-off with significantly different variances, a Games-Howell analysis, which is a post hoc test, is performed.

Table 6. Games-Howell test, Multiple comparisons

Dependent variable: Games-Howell	Rating	Difference of means (I-J)	Error	Sig.	95% confidence interval		
					Lower limit	Upper limit	
(I) Cutting	1	2	-1,1659*	0,2299	0,000	-1,714	-0,618
		3	-1,6386*	0,1670	0,000	-2,043	-1,235
	2	1	1,1659*	0,2299	0,000	0,618	1,714
		3	-,4727*	0,1680	0,019	-0,879	-0,067
	3	1	1,6386*	0,1670	0,000	1,235	2,043
		2	,4727*	0,1680	0,019	0,067	0,879

*. The difference in means is significant at the 0.05 level.

The Games-Howell test confirms significant differences between the three cuts evaluated with the written test. Among the results, it is found that in Cut 1, the mean is significantly lower than in Cut 2 and Cut 3, and in the case of Cut 2, it also has a considerably lower mean than in Cut 3. Additionally, it can be determined that the ratings increase consistently from Cut 1 to Cut 3, and all comparisons show statistically significant differences at the 5% level.

The behavior of the results in the face-to-face evaluation indicates that a greater dispersion of the results is generated in the assessment with a face-to-face written evaluation method.

Table 7. Comparison with other studies on comparative studies between student assessment models

Ref	Target audience	Findings
[4]	Undergraduate students at a public university in Bosnia	The results showed a significant difference in student performance, and that it depends on the learning environment. In particular, students who attended online classes.
[6]	University students enrolled in Spanish universities during the 2017/2018 academic year.	Virtual modalities favor a greater interdependence of learning strategies, study time, and motivation, which could reduce stress and favor academic performance.
[8]	University students of the Education Degrees, belonging to the Faculties of Humanities and Sciences of the Education of Andalusia (Spain)	presents the need to implement cross-training as a digital learning strategy to strengthen adaptive processes and personal and academic performance.
[20]	Bibliometric Review	It highlights the multidisciplinary nature of YouTube as an educational medium.

According to the table above, similar findings are presented with the studies related to strengthening students' performance at the university level, comparing the learning and evaluation in person with the virtual, online, or digital process. The results showed a difference in the behavior of grades and academic performance, linked to the learning environment. Likewise, studies show that the grade obtained depends on the actual effort of students. The study conducted at the Bosnian Public University agrees with this research because it was concluded that students perform better and their grades are stable in digital environments such as social networks. The study conducted by [20] is the only one that emphasizes the work of a social network as a mechanism to strengthen students' academic performance. The study confirms that presenting deliverables in media such as social networks limits the dispersion in grades and improves performance in each cut.

However, linear regressions can be performed to estimate how much the type of evaluation influences the grade and to see if this relationship is statistically significant. Additionally, a study can be conducted with a sample of different academic semesters and classify the ages and careers. In summary, all studies aim to define student success by comparing two learning environments (digital and face-to-face); no study performs an analysis relating tests or exams to lectures.

4. Conclusions

Among the main findings of the study analyzing two evaluation modalities with differentiated learning environments are results that suggest that the evaluation mediated by Facebook tends to promote greater homogeneity in the grades received by students, thus strengthening their student performance, probably due to collaborative work and previous feedback from classmates. This positively guarantees participation and understanding of conceptual knowledge. On the other hand, with greater acceptance in the academic environment, the written test evidenced a significant dispersion in the results in the academic cuts.

Under academic rigor, these results point to enriching knowledge measurement strategies, incorporating digital learning environments and Web 2.0 as spaces for exchanging knowledge and evaluating the students' academic competencies. Similarly, it is possible to reflect on the effectiveness of written tests, which may not respond to the qualities of the university student's profile in a changing environment.

5. Limitations

The study was limited to a single university, and the results are based on a specific group of students who used Facebook as an assessment tool. At the same time, the rest of the institution conducted their assessments face-to-face. Therefore, the findings may not be generalizable to other educational contexts. In addition, social networks are constantly evolving, so the results must be interpreted in the study's time frame.

Declaration of competing interest

The authors declare that they have no known financial or non-financial competing interests in any material discussed in this paper.

Funding information

No funding was received from any financial organization to conduct this research.

References

- [1] S. Salazar-palomino, "Digital competences and collaborative skills among university students," vol. 7, no. 1, pp. 1–18, 2025.
- [2] L. Tašaković and Ö. Büyükdağlı, "Impact assessment methods for teaching activities on sustainable development goals in higher education institutions: A case study from a Bosnian university," *Heritage and Sustainable Development*, vol. 6, no. 2, pp. 445–458, 2024, doi: 10.37868/hsd.v6i2.386.
- [3] N. N. Nayef and L. M. Bilani, "The role of knowledge management in improving the quality of projects: Research from Iraq," *Heritage and Sustainable Development*, vol. 6, no. 1, pp. 145–156, 2024, doi: 10.37868/hsd.v6i1.391.
- [4] S. Sarač and B. Duraković, "Analysis of student performances in online and face-to-face learning: A case study from a Bosnian public university," *Heritage and Sustainable Development*, vol. 4, no. 2, pp. 87–94, 2022, doi: 10.37868/HSD.V4I2.91.
- [5] O. Lefebvre, "Modeling innovation and regulation thanks to game theory: Bertrand competition," *Heritage and Sustainable Development*, vol. 4, no. 2, pp. 165–173, 2022, doi: 10.37868/hsd.v4i2.150.
- [6] R. Chacón-Cuberos, M. Pérez-Mármol, E. García-Mármol, E. M. Olmedo-Moreno, J. Expósito-López, and M. Castro-Sánchez, "El papel de la virtualidad en las estrategias de aprendizaje y la prevención del estrés en educación superior," *Educación XXI*, vol. 27, no. 2, pp. 89–113, 2024, doi: 10.5944/educxx1.35855.
- [7] J. Romeo, D. Ramírez, J. Antonio, and H. Martínez, "Evaluación formativa y calidad de la retroalimentación : diseño y validación de escalas para docentes escolares," vol. 27, no. 2, pp. 167–194, 2024.
- [8] Ó. Gavín-Chocano, I. García-Martínez, E. Pérez-Navío, and A. Luque de la Rosa, "Learner engagement, motivación académica y estrategias de aprendizaje de estudiantes universitarios," *Educación XXI*, vol. 27, no. 1, pp. 57–79, 2024, doi: 10.5944/educxx1.36951.
- [9] L. Aperribai and L. Cortabarria, "Barreras para la creatividad personal en estudiantes universitarios españoles y portugueses," *Educación XXI*, vol. 27, no. 1, pp. 81–104, 2024.
- [10] M. S. Ibarra-Sáiz, G. Rodríguez-Gómez, J. F. Lukas-Mujika, and A. Santos-Berrondo, "Methods and instruments to assess learning outcomes in master's degrees. Analysis of teachers' perception of their evaluative practice," *Educacion XXI*, vol. 26, no. 1, pp. 21–45, 2023, doi: 10.5944/educxx1.33443.
- [11] J. Fernandez-Río, J. A. Cecchini, J. Lopes, H. Silva, and Â. Leite, "Autoeficacia, autorregulación y aprendizaje cooperativo en estudiantes españoles y portugueses de Educación Secundaria," *Educación XXI*, vol. 26, no. 1, pp. 117–139, 2023, doi: 10.5944/educxx1.33339.
- [12] R. García-Ruiz, M. Buenestado-Fernández, and M. S. Ramírez-Montoya, "Assessment of Digital Teaching Competence: Instruments, results and proposals. Systematic literature review," *Educacion XXI*, vol. 26, no. 1, pp. 273–301, 2023, doi: 10.5944/educxx1.33520.
- [13] E. Panadero, J. Fraile, and D. G. Pérez, "Transición a educación superior y evaluación: un estudio longitudinal anual," *Educacion XXI*, vol. 25, no. 2, pp. 15–37, 2022.
- [14] L. Parcerisa, J. Jacovkis, C. Lindín, and I. Màrquez, "Digital Sovereignty and Education: A Missing Link in the Literature," *REICE. Revista Iberoamericana Sobre Calidad, Eficacia y Cambio en Educacion*, vol. 22, no. 2, pp. 151–168, 2024, doi: 10.15366/reice2024.22.2.009.
- [15] P. Rivera-Vargas, B. Mateu-Luján, S. Rappoport, and Y. Gamboa, "Digitization of Schools and the Use of Mobiles in the Classroom. Analysis of the Spanish Case," *REICE. Revista Iberoamericana Sobre Calidad, Eficacia y Cambio en Educacion*, vol. 21, no. 4, pp. 25–43, 2023, doi: 10.15366/reice2023.21.4.002.

- [16] C. Monge and P. Gómez, “Differences in Innovation Factors in Science and Mathematics Teachers in CSE,” *REICE. Revista Iberoamericana Sobre Calidad, Eficacia y Cambio en Educacion*, vol. 21, no. 4, pp. 87–109, 2023, doi: 10.15366/reice2023.21.4.005.
- [17] E. Practices and P. Renewal, “PRESENTACIÓN Prácticas Educativas de Renovación Pedagógica en la Actualidad : Una Perspectiva Crítica,” vol. 21, no. 2, pp. 5–8, 2023.
- [18] M. del M. Sánchez Vera and P. López Vicent, “Movimientos de Renovación Pedagógica en Redes con Perspectiva Ecosocial: El Caso de Teachers for Future,” *REICE. Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, vol. 21, no. 2, pp. 51–69, 2023, doi: 10.15366/reice2023.21.2.003.
- [19] P. J. Carrillo-López and D. Hortigüela-Alcalá, “Teacher Formative Evaluative Practices and Academic Performance in Primary School Students,” *REICE. Revista Iberoamericana Sobre Calidad, Eficacia y Cambio en Educacion*, vol. 21, no. 1, pp. 5–20, 2023, doi: 10.15366/reice2023.21.1.001.
- [20] M. P. Colás-Bravo and I. Quintero-Rodríguez, “YouTube and Learning: A Systematic Literature Review,” *REICE. Revista Iberoamericana Sobre Calidad, Eficacia y Cambio en Educacion*, vol. 21, no. 1, pp. 47–63, 2023, doi: 10.15366/reice2023.21.1.003.