STATISTICAL QUALITY CONTROL OF ENGINEERING LICENSURE EXAMINATION PERFORMANCE: P-CHART, ATTRIBUTE CAPABILITY, AND ISHIKAWA ANALYSIS

Nicanor F. Branzuela Jr.
University of Science and Technology of Southern Philippines
C.M. Recto Avenue, Lapasan, Cagayan de Oro City 9000, Philippines
Correspondence: nicanor.branzuela@ustp.edu.ph
Date of Publication: July 30, 2026

ABSTRACT*- This retrospective program-level study evaluated engineering licensure examination performance through an integrated statistical quality-control framework. Aggregated institutional records covered 1,069 examinees from Civil Engineering, Electrical Engineering, and Mechanical Engineering across 17 available examination periods from 2022 to 2025. Non-passers were treated as non-conforming outcomes. Variable-subgroup p-charts assessed statistical stability, while an attribute capability assessment compared each program's pooled non-conformance rate with a predefined 40% institutional action threshold. A literature-informed Ishikawa diagram organized probable causes for subsequent validation. Civil Engineering was statistically stable but not capable: 346 of 701 examinees did not pass (49.36%), and no point exceeded its control limits. Electrical Engineering was capable overall (25 of 153; 16.34%) but not fully stable because April 2022 and August 2024 exceeded their upper control limits. Mechanical Engineering showed the strongest overall performance (28 of 215; 13.02%; yield = 86.98%), although the August 2023 subgroup contained only one examinee and therefore had limited interpretive value. The findings demonstrate that passing percentage alone cannot distinguish predictable underperformance from isolated special-cause variation. Combining p-charts, threshold-based attribute capability assessment, and structured root-cause mapping provides a more actionable basis for program review. Capability classifications remain conditional on the selected institutional threshold, and the proposed Ishikawa causes require validation using cohort records, interviews, and subject-level performance data.*

Keywords- Attribute data, Engineering education, Ishikawa diagram, licensure examination, P-chart, Statistical qualitycontrol

1. INTRODUCTION

Licensure examination results are among the most visible indicators of professional-program outcomes. They influence curriculum review, student-support planning, public accountability, and stakeholder judgments about graduate readiness. Yet institutional reports commonly summarize performance only as passing percentages. A single percentage cannot show whether an observed result reflects routine variation, an isolated disturbance, or a persistently inadequate process level.

Contemporary quality-assurance research therefore favors multidimensional and evidence-based evaluation. Educational quality is shaped by interacting instructional, institutional, learner, and resource conditions [1], while higher-education quality assurance increasingly emphasizes systematic evaluation and improvement rather than compliance alone [2]. Statistical process control (SPC) extends this logic by examining performance as a process observed repeatedly over time. Educational applications have shown that control charts can identify unusual periods that ordinary summaries may conceal [3-5].

A p-chart is appropriate when outcomes are binary and subgroup sizes vary. In the present context, each examinee was classified as conforming (passed) or non-conforming (did not pass), and each examination period formed a subgroup. The chart estimates a pooled non-conformance rate and subgroup-specific three-sigma limits, thereby separating common-cause variation from potential special-cause signals [6-9]. However, statistical stability is not equivalent to acceptable performance. A program may be stable at an undesirable non-conformance level, or it may have a favorable overall rate while containing isolated unstable periods.

Attribute capability assessment addresses this distinction by comparing observed non-conformance with a defined requirement. Capability concepts link process performance to specifications, yield, and non-conformities [10-12]. For binary educational outcomes, the comparison must be interpreted cautiously because standard Cp and Cpk indices are designed for continuous measurements. This study therefore uses a threshold-based attribute assessment and explicitly labels its capability ratio as a study-defined comparison rather than a conventional process capability index.

Statistical signals also require diagnostic follow-up. Fishbone or Ishikawa diagrams organize plausible causes into categories and support systematic investigation [13-16]. They do not establish causality, but they help translate an abstract signal into evidence to be collected from learners, faculty, program records, instructional materials, and data systems.

The study integrated these three functions-stability monitoring, threshold-based capability assessment, and root-cause mapping-to evaluate Civil Engineering, Electrical Engineering, and Mechanical Engineering licensure examination performance. It aimed to determine each program's pooled non-conformance rate, identify p-chart signals, classify capability relative to a 40% institutional action threshold, and specify evidence priorities for continuous improvement.

METHODOLOGY

Research design and data source

A retrospective descriptive-evaluative design was used. The dataset consisted of all available aggregated institutional licensure records for the three engineering programs from 2022 to 2025. For every examination period, the records reported the total number of takers, passers, and non-passers.

The analysis included five Civil Engineering periods, six Electrical Engineering periods, and six Mechanical Engineering periods. Individual names, student numbers, and other personal identifiers were neither required nor reported.

Each examination period was treated as a subgroup. The total number of takers was denoted by n(i), the number of non-passers by d(i), and the subgroup non-conformance proportion by p(i) = d(i)/n(i). The pooled center line for each program was calculated as p-bar = sum[d(i)]/sum[n(i)].

P-chart analysis

Subgroup sizes differed, the upper and lower control limits were computed separately for each period: UCL(i) = p-bar + 3{p-bar(1-p-bar)/n(i)}^(1/2) and LCL(i) = max[0, p-bar - 3{p-bar(1-p-bar)/n(i)}^(1/2)]. A point above the UCL was interpreted as an unfavorable special-cause signal requiring investigation.

Threshold-based attribute capability assessment

Program-level attribute performance was summarized using pooled non-conformance, yield (1 - p-bar), non-conforming outcomes per million observations (p-bar x 1,000,000), Z.Bench = Phi^(-1)(yield), and an illustrative long-term sigma conversion (Z.Bench + 1.5). A study-defined threshold performance ratio was calculated as TPR = 0.40/p-bar. TPR values above 1.00 indicate that the observed non-conformance rate is below the 40% action threshold. These metrics are descriptive transformations of binary outcomes; TPR is not Cp, Cpk, Pp, or Ppk. A program was classified as capable only in relation to the predefined threshold.

Ishikawa analysis

The effect placed at the head of the Ishikawa diagram was high or unstable non-conformance in engineering licensure performance. Probable causes were organized under learner preparedness, instructional process, program/system, faculty and support, materials and resources, and measurement/data. The categories were derived from the statistical patterns and quality-improvement literature.

3. RESULTS AND DISCUSSION

Overall program conditions

The three programs represented distinct quality conditions (Table 1). Civil Engineering had the largest number of examinees and the highest pooled non-conformance rate. Electrical Engineering had a comparatively low pooled rate but two above-UCL periods. Mechanical Engineering had the lowest pooled rate and no above-UCL points, although one subgroup contained only one examinee.

Table 1: Overall p-chart and threshold-based attribute capability summary

Program	Takers	Non-passers	p-bar	Yield	Signals	TPR	Quality condition
Civil Engineering	701	346	49.36%	50.64%	0	0.81	Stable; not capable
Electrical Engineering	153	25	16.34%	83.66%	2	2.45	Capable; unstable
Mechanical Engineering	215	28	13.02%	86.98%	0	3.07	Capable; mostly stable

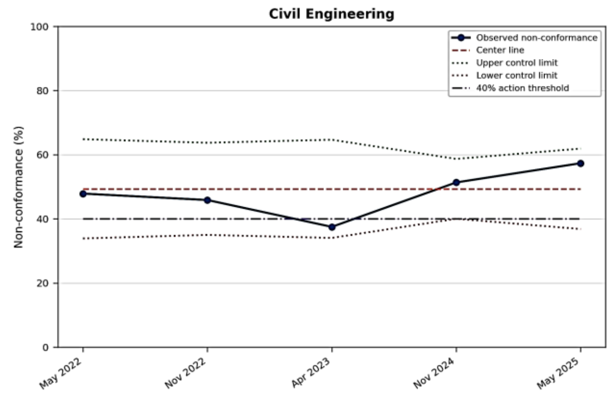

Figure 1. Variable-subgroup p-chart for Civil Engineering.

Civil Engineering recorded 346 non-passers among 701 examinees (49.36%). All five points remained within their subgroup-specific limits, including the highest observed rate in May 2025 (57.34%). The chart therefore suggests common-cause variation around a consistently high center. Because the pooled rate exceeded the 40% threshold (TPR = 0.81), the program is stable but not capable. The priority is not to search for a single anomalous cohort but to review recurring system conditions, including curriculum alignment, competency remediation, mock-board preparation, review attendance, and student readiness.

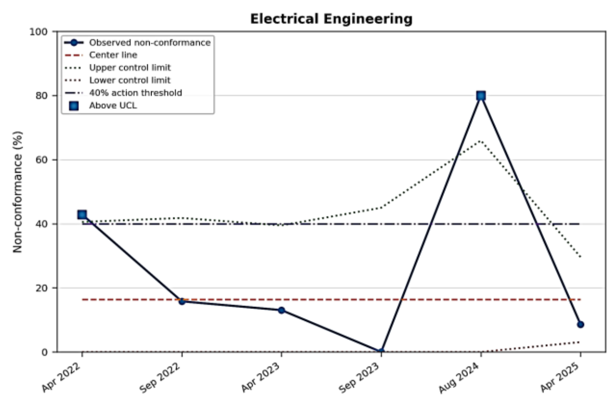

Figure 2. Variable-subgroup p-chart for Electrical Engineering.

Electrical Engineering recorded 25 non-passers among 153 examinees (16.34%), corresponding to a yield of 83.66% and TPR of 2.45. April 2022 (42.86%) and August 2024 (80.00%) exceeded their respective UCLs. The program is therefore capable in the threshold sense but not fully stable. The August 2024 subgroup contained only five examinees, so the signal should be investigated without overgeneralization. Useful evidence includes cohort composition, subject-level scores, review implementation logs, attendance, faculty assignments, and any changes in examination or reporting practice.

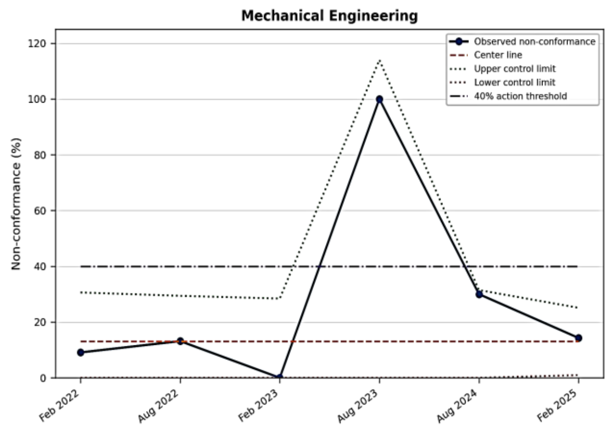

Figure 3. Variable-subgroup p-chart for Mechanical Engineering.

Mechanical Engineering recorded 28 non-passers among 215 examinees (13.02%), the lowest pooled rate and the highest yield (86.98%) among the programs. Its TPR of 3.07 indicates the largest margin below the action threshold. No point exceeded the UCL. Nevertheless, August 2023 produced 100% non-conformance because the subgroup consisted of only one examinee. The corresponding UCL was extremely wide, illustrating how very small n can make a p-chart mathematically valid but practically uninformative. Continued monitoring should therefore include a minimum-subgroup interpretation rule and rolling aggregation when cohorts are exceptionally small.

Attribute performance interpretation

The attribute indicators reinforced the distinction between stability and adequacy. Civil Engineering had an estimated Z.Bench of 0.02 and illustrative long-term sigma level of 1.52, whereas Electrical Engineering had 0.98 and 2.48, and Mechanical Engineering had 1.13 and 2.63, respectively.

The threshold itself is consequential. A 40% non-conformance action level classifies Electrical and Mechanical Engineering as capable and Civil Engineering as not capable, but another evidence-based threshold could change these classifications. Future institutional use should justify the threshold through policy, historical performance, accreditation expectations, or comparison with national and peer-institution benchmarks. Statistical control limits and performance thresholds answer different questions and should not be substituted for one another.

Ishikawa-guided diagnostic priorities

Figure 4 converts the statistical findings into an investigation structure. The diagram does not claim that the listed factors caused failure. Rather, it identifies categories under which the institution can collect evidence. The program-specific logic is straightforward: stable but high non-conformance points toward recurring system and instructional conditions; isolated above-UCL points point toward period-specific changes; and very small subgroups point toward data-quality and interpretation safeguards.

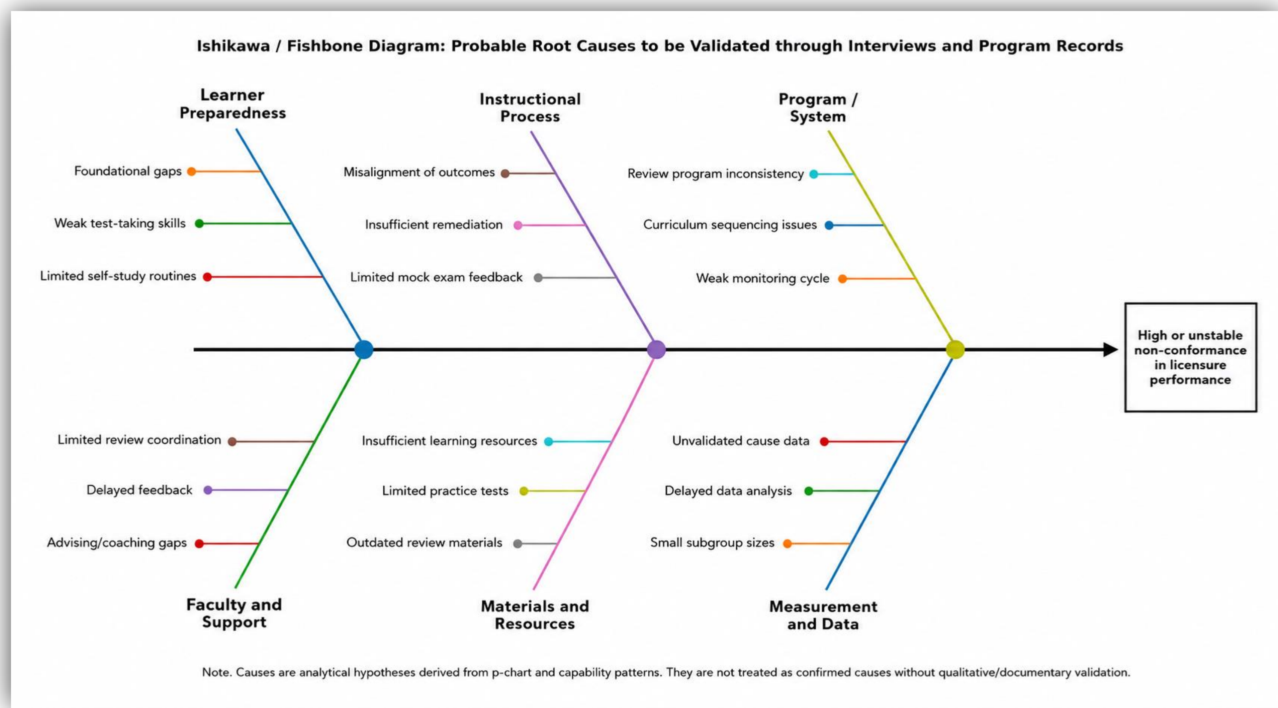

Figure 4. Ishikawa diagram of probable causes requiring validation.

Table 2: Program-specific diagnostic focus and evidence for validation

Program	Statistical condition	Primary diagnostic focus	Suggested evidence
Civil Engineering	Stable but above threshold	Program/system; instructional process; learner preparedness	Curriculum mapping; subject-level board ratings; mock-board results; review attendance; remediation records; faculty and student interviews.
Electrical Engineering	Below threshold with two above-UCL periods	Period-specific instructional, cohort, support, and data factors	Cohort profiles; class-size notes; faculty assignments; review logs; attendance; changes in exam preparation or data coding.
Mechanical Engineering	Below threshold and mostly stable; one n = 1 subgroup	Measurement/data quality and monitoring rules	Minimum-subgroup rule; rolling aggregation; source-record validation; continued Phase II monitoring.

LIMITATIONS

The analysis was limited to one institution, three programs, and 17 examination periods. Each program had only five or six subgroups, so the estimated control limits are provisional and sensitive to the available cohorts. Subgroup sizes were unequal, and one Mechanical Engineering period had n = 1. The study used aggregated results and therefore could not examine first-time-taker status, academic history, subject-level board ratings, review attendance, demographic factors, or other predictors. The 40% threshold was an institutional action value rather than a national specification. Finally, the Ishikawa categories were not empirically validated and cannot support causal conclusions.

CONCLUSION AND RECOMMENDATIONS

Engineering licensure performance is more informative when evaluated as a repeated quality process rather than as a sequence of isolated passing percentages. The combined framework distinguished three conditions: Civil Engineering was stable but not capable relative to the 40% action threshold; Electrical Engineering was capable overall but unstable because of two special-cause signals; and Mechanical Engineering was capable and mostly stable but contained one practically uninterpretable one-person subgroup.

The institution should use program-specific responses. Civil Engineering requires sustained system-level improvement rather than a one-period corrective action. Electrical Engineering requires targeted investigation of April 2022 and August 2024 before its favorable pooled capability is treated as dependable. Mechanical Engineering should preserve effective practices while strengthening small-cohort interpretation and data validation. Across programs, the institution should establish a verified Phase I baseline, freeze the limits for prospective Phase II monitoring, and re-baseline only after documented process change.

Future analyses may incorporate national passing rates, institutional targets, first-time and repeat examinee status, subject-level ratings, academic records, and review-program participation. Interviews and document analysis should be used to validate the Ishikawa hypotheses. A Plan-Do-Check-Act cycle can then connect each examination result to targeted intervention, prospective monitoring, and evidence-based standardization.

DATA AND ETHICAL STATEMENT

The study analyzed aggregated, de-identified institutional records and involved no direct interaction with individual examinees. Names, student numbers, and other personally identifiable information were not included in the analytical dataset or reported in the manuscript.

REFERENCES

[1] Rodríguez, J. V., Neira Rodado, D., Crissien Borrero, T. and Parody, A., "Multidimensional indicator to measure quality in education," International Journal of Educational Development, 89: 102541 (2022). https://doi.org/10.1016/j.ijedudev.2021.102541

[2] Khuram, S., Rehman, C. A., Nasir, N. and Elahi, N. S., "A bibliometric analysis of quality assurance in higher education institutions: Implications for assessing university societal impact," Evaluation and Program Planning, 99: 102319 (2023). https://doi.org/10.1016/j.evalprogplan.2023.102319

[3] Gessa, A., Marín, E. and Sancha, P., "A practical application of statistical process control to evaluate the performance rate of academic programmes: Implications and suggestions," Quality Assurance in Education, 30(4): 571-588 (2022). https://doi.org/10.1108/QAE-03-2022-0065

[4] Masnar, H. M. and Namoco, C. S. Jr., "Utilizing student performance data in monitoring new curriculum in construction technology using statistical process control (SPC)," Science International (Lahore), 36(1): 59-63 (2024).

[5] Capilitan, F. T. Jr., Duero, J. G. Q., Daleon, J. D., Dumaog, A. C. and Namoco, C. S. Jr., "Assessment on the performance of teacher education graduates in the Licensure Examination for Teachers using statistical quality control: A case study of University of Science and Technology of Southern Philippines-Cagayan de Oro," Science International (Lahore), 34(3): 327-330 (2022).

[6] Joekes, S. and Barbosa, E. P., "An improved attribute control chart for monitoring non-conforming proportion in high quality processes," Control Engineering Practice, 21(4): 407-412 (2013). https://doi.org/10.1016/j.conengprac.2012.12.005

[7] Wang, H., "Comparison of p control charts for low defective rate," Computational Statistics & Data Analysis, 53(12): 4210-4220 (2009). https://doi.org/10.1016/j.csda.2009.05.024

[8] Montgomery, D. C., Introduction to Statistical Quality Control, 8th ed., Wiley, Hoboken, New Jersey (2020).

[9] Woodall, W. H., "Controversies and contradictions in statistical process control," Journal of Quality Technology, 32(4): 341-378 (2000). https://doi.org/10.1080/00224065.2000.11980013

[10] Wu, C.-W., Pearn, W. L. and Kotz, S., "An overview of theory and practice on process capability indices for quality assurance," International Journal of Production Economics, 117(2): 338-359 (2009). https://doi.org/10.1016/j.ijpe.2008.11.008

[11] Wang, H., Yang, J. and Hao, S., "Two inverse normalizing transformation methods for the process capability analysis of non-normal process data," Computers & Industrial Engineering, 102: 88-98 (2016). https://doi.org/10.1016/j.cie.2016.10.014

[12] Almeida, M. C., Oliveira, L. G., Rotella Junior, P. and Peruchi, R. S., "Multivariate process capability analysis with decision-maker preferences," Computers & Industrial Engineering, 198: 110664 (2024). https://doi.org/10.1016/j.cie.2024.110664

[13] Johannssen, A., Chukhrova, N. and Castagliola, P., "The performance of the hypergeometric np chart with estimated parameter," European Journal of Operational Research, 296(3): 873-899 (2022). https://doi.org/10.1016/j.ejor.2021.06.056

[14] Cox, M. and Sandberg, K., "Modeling causal relationships in quality improvement," Current Problems in Pediatric and Adolescent Health Care, 48(7): 182-185 (2018). https://doi.org/10.1016/j.cppeds.2018.08.011

[15] Carvalho, R., Lobo, M., Oliveira, M., Oliveira, A. R., Lopes, F., Souza, J., Ramalho, A., Viana, J., Alonso, V., Caballero, I., Santos, J. V. and Freitas, A., "Analysis of root causes of problems affecting the quality of hospital administrative data: A systematic review and Ishikawa diagram," International Journal of Medical Informatics, 156: 104584 (2021). https://doi.org/10.1016/j.ijmedinf.2021.104584

[16] Jipos, A. M. C., Jamito, C. P., Camelotes, J. B., Baguio, R. H. and Namoco, C. S. Jr., "Root-cause analysis of recurring flour packaging printing defects utilizing quality control tools," Science International (Lahore), 35(6): 725-728 (2023).