Journal of Technological Development 9[22]1-13: e30922113

Article

Learning analytics with CAD traces and predictive models: Reliability and performance

Analítica de aprendizaje con trazas CAD y modelos predictivos: Confiabilidad y desempeño

Corral-Verdugo, Alex*[a], Sepúlveda-Romo, Adrián[b], Jimenez-Lopez, Eusebio[c] and León-Rochin, German[d]

[a] Universidad Tecnologica del sur de Sonora / Universidad La Salle Noroeste • ODI-7910-2025 • 0009-0002-7645-9857 • 072915

[b] Universidad Tecnologica del sur de Sonora • OJT-1262-2025 • 0000-0002-7639-3512 • 2072915

[c] Instituto Tecnológico Superior de Cajeme / Universidad La Salle Noroeste • 0000-0001-6893-3550 • 66512

[d] Universidad Tecnologica del sur de Sonora • OQK-6272-2025 • 0000-0002-0120-7922 • 1265462

Classification:

Area: Social Sciences and Humanities
Field: Education
Discipline: Education
Subdiscipline: Comparative Education

https://doi.org/10.35429/JTD.2025.9.22.3.1.13

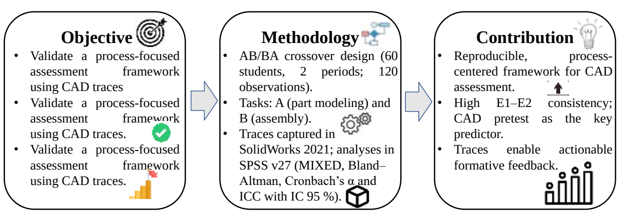
History of the article:
Received: January 13, 2025
Accepted: November 10, 2025

Check for updates

* ✉ [acorral@uts.edu.mx]

Abstract

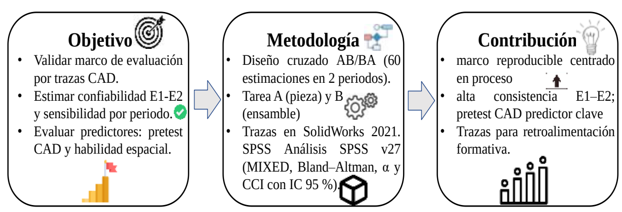
This study gives us a proposal and validation of a trace-based evaluative framework for CAD courses in technical education. It integrates entry predictors (CAD pretest and spatial ability), process logs (operations, errors, undo actions, latencies and fully defined sketches), and rubric-based scoring to offer instruction-sensitive formative feedback. With 60 students under a crossed AB/BA design (120 observations), we estimated inter-rater reliability and predictive models. Consistency was high (α = .915; ICC = .849/.919) and there was no systematic bias between evaluations ($t(23) = -0.02$, p = .986; Bland–Altman). Performance was explained mainly by the CAD pretest, with additional contribution from spatial ability; E1 and E2 correlated strongly (r = .85). The study showed sequence-differential attrition (p = .006) which did not present selection bias. The study offers a replicable protocol and metrics that can be used by any instructor, promoting more efficient and equitable assessment decision-making.



CAD learning analytics; process traces

Resumen

El presente estudio describe una propuesta y validación de un marco evaluativo basado en trazas para cursos de CAD (Computer Aided Desing) en educación técnica. Integra predictores de entrada (pretest de CAD y visión espacial), registros de proceso (operaciones, errores, deshacer y latencias, bocetos definidos) y calificación por rúbrica para ofrecer retroalimentación formativa sensible a la instrucción. Con 60 estudiantes bajo un diseño cruzado AB/BA (120 observaciones), se estimó la confiabilidad interevaluador y los modelos predictivos. La consistencia fue alta (α=.915; ICC=.849/.919) y no hubo sesgo sistemático entre evaluaciones ($t(23)=-0.02$, p=.986; Bland–Altman). El desempeño se explicó principalmente por el pretest de CAD, con aporte adicional de la visión espacial; E1 y E2 se correlacionaron fuertemente (r=.85). En el estudio existió atrición diferencial por secuencia (p=.006) la cual no presento sesgo de selección. El estudio ofrece un protocolo replicable y métricas las cuales pueden ser utilizadas por cualquier docente, promoviendo toma de decisiones evaluativas más eficientes y equitativas.



CAD, analítica del aprendizaje, trazas de proceso

Area: Development of strategic leading-edge technologies and open innovation for social transformation.

Citation: Corral-Verdugo, Alex, Sepúlveda-Romo, Adrián, Jimenez-Lopez, Eusebio and León-Rochin, German. [2025]. Development of strategic leading-edge technologies and open innovation for social transformation. Journal of Technological Development. 9 [22] 1-13: e30922113

Introduction

In recent years, computer-aided design (CAD) has served as a central element in technical and engineering training. This enables geometric representation, articulation and collaboration, iteration and feedback throughout the entire creation and design process. Therefore, it is important to focus not only on the final product, but also on the process that students follow when creating their models (Bojčetić *et al.*, 2021; Contero *et al*., 2023; Xie *et a*l., 2014).

In large populations, manual assessment becomes costly and diverse, which is why there are proposals for automated assessment and learning analytics approaches that provide in-depth, timely and effective feedback (Eltaief *et al*., 2024; Jaakma and Kiviluoma, 2019; Company *et al*., 2015).

Approaching an assessment process using a 3D model through learning analytics provides great value that is applicable both theoretically and methodologically. In this way, it converges as a scaffolding that is capable of interpreting process data and closing the feedback loop through learning dashboards and continuous improvement frameworks focused on learning, provided that instructional design guides capture, interpretation, and action (Wang, 2021; Sailer *et al*., 2024; Paulsen and Lindsay, 2024; Susnjak *et al*., 2022; Anghel *et al*., 2024). Applying all these principles to the CAD context is natural when performing sequences of sketching, editing, commands such as undo/redo, and design regenerations.

It is time-based analytics with a design focus that allows difficulties to be anticipated and specific support to be provided at the most opportune moment (Wang, 2021; Sailer *et al*., 2024).

Speaking specifically about the field of CAD design, empirical evidence suggests that traces are related to the instruction given. Some pedagogical adjustments provided outside the design software leave certain detectable traces or footprints in action patterns within the tool used, and all this enables estimation during the process and not only at the end of it (Xie *et al*., 2014).

Similarly, progress has been made in proposals for visual analytics and formative feedback that explore the structure of the entire record of the parametric models created, in order to detect all the deficiencies that arose during the construction of the model, with a view to guiding more timely and accurate pedagogical interventions (Otto & Mandorli 2021; Mandorli, & Otto, 2024; Otto & Mandorli, 2025). This is why the evaluation process moves from individual geometric verification to an explicit assessment of the strategy for constructing and modifying the model to be created, in addition to dimensions that can be inferred from the traces and history of all the parametric operations performed (Contero *et al*., 2023; Aranburu *et al*., 2023; Zou *et al*., 2023).

Several recent studies suggest lines of research that help strengthen the evaluation process. Firstly, the application of self-assessment tools, which provide verification of both geometric accuracy and design intent, offers promising results in higher-level courses by allowing the robustness of the applied model to be estimated in the face of parametric changes and reducing the workload for the teacher (Jaakma and Kiviluoma, 2019).

Secondly, the application of coordinated rubrics that formalise quality criteria for all parametric modelling carried out, as well as generating a logical order of operations used, in addition to the complete definition of sketches, nomenclature and the current capacity for modification in designs, which places teaching alongside assessment and provides greater consistency between different assessors (Company *et al*., 2015).

Thirdly, there are adjustable multi-criteria tools that weigh geometric dimensions, based on operations and parameters in accordance with very specific teaching objectives, which present recent improvements in both precision and flexibility (Eltaief *et al*., 2024; Nobes, 2025). These lines aim to scale the grade obtained, reduce any type of bias and, above all, offer an approach to assessment and how it is modelled in student practice (Bojčetić *et al*., 2021).

It is necessary to consider that the performance shown by students in CAD is not solely a consequence of their performance with this tool.

ISSN 2531-2960
RENIECYT: 1702902

Corral-Verdugo, Alex, Sepúlveda-Romo, Adrián, Jimenez-Lopez, Eusebio and León-Rochin, German. [2025]. Development of strategic leading-edge technologies and open innovation for social transformation. Journal of Technological Development. 9 [22] 1-13: e30922113
https://doi.org/10.35429/JTD.2025.9.22.3.1.13

Measurement using individual predictors is key to correctly interpreting the results. Similarly, prior knowledge shows high stability between measurements and its relationship with learning, although this relationship varies depending on the context. It also functions as a fundamental covariate when analysing instructional effects (Simonsmeier *et al*., 2022). Similarly, spatial ability is associated with performance in design graphics and CAD courses and can be developed through specific instruction.

Recent reviews on the evaluation of technical sketches and the use of immersive environments support the relevance of considering these variables in course design (Bartlett & Camba, 2023; Dilling & Vogler, 2021; Gittinger & Wiesche, 2024; Merzdorf *et al*., 2024; Zhu *et al*., 2023). Consequently, the integration of these predictors with CAD traces allows us to explain the disparity that still exists depending on the different learning paths adopted, enabling us to focus timely support where it is needed (Dilling & Vogler, 2021; Merzdorf *et al*., 2024; Riestra-González *et al*., 2021; Sailer *et al*., 2024; Simonsmeier *et al*., 2022; Wang, 2021). Even with all this, there are significant gaps in real classroom environments. Despite the various solutions that exist, these prioritise product metrics such as geometric matching, operation counting and, more recently, the incorporation of more flexible forms such as construction and modification capacity as weighted criteria (Eltaief *et al*., 2024; Company *et al*., 2015).

However, there are very few studies that combine a quasi-experimental classroom design with task counterbalancing, analysis of process traces, and the management and use of rubrics as an assessment process (Otto & Mandorli, 2021; Mandorli, & Otto, 2024; Nobes, 2025). It is also worth mentioning the lack of evidence in upper secondary technical education, due to the large cohorts involved, time constraints, and the need for reproducible, scalable, and economical methods (Bojčetić *et al*., 2021).

This work evaluates the feasibility of using CAD software traces as an assessment method in SolidWorks part construction and assembly tasks, integrating three components: 1) process metrics per session, 2) product ratings, and 3) individual predictors.

Input predictors such as CAD pretest, spatial ability, process traces, and product performance are added to facilitate more comprehensive formative feedback (see Figure 1). In particular, operations per minute, milliseconds per operation, proportion of fully defined sketches, undo options per minute, errors per minute, and total time are analysed. Two product evaluations at different times (E1 and E2) were considered, and the CAD pretest and spatial ability were incorporated as covariates.

This was achieved through a cross-over design with counterbalancing by task. In addition, we studied whether the traces capture instructional differences between periods, the consistency and absence of bias in the evaluation, and the additional explanatory power of the predictors on the performance shown.

This provides three elements: 1) a replicable framework for implementing CAD courses with trace-based assessment, 2) evidence on which process metrics are most informative for formative feedback, and 3) guidelines for integrating rubrics, assisted self-assessment, and learning analytics dashboards into CAD, all of this in line with closed-loop frameworks considering the instructional sensitivity offered by traces (Sailer *et al*., 2024; Paulsen and Lindsay, 2024; Susnjak *et al*., 2022; Xie *et al*., 2014).

This study is highly relevant because it is firmly rooted in Mexican technical education, where the curricula of institutions such as DGETI and CONALEP incorporate computer-aided design, modelling and drawing (CAD) modules in mechatronics and related areas. That is why CAD skills form the basis of a key curricular competency in student training (SEMS–COSFAC, n.d.; CONALEP, 2023).

In the productive sphere, there is great external relevance, as this is supported by the demand for talent in the country's manufacturing and aerospace sectors, with constant contributions to employment and production, thus creating an evident need for human capital specialised in these technical areas that are necessary in the country (INEGI, 2025; Secretaría de Economía, 2020).

ISSN 2531-2960
RENIECYT: 1702902

Box 1

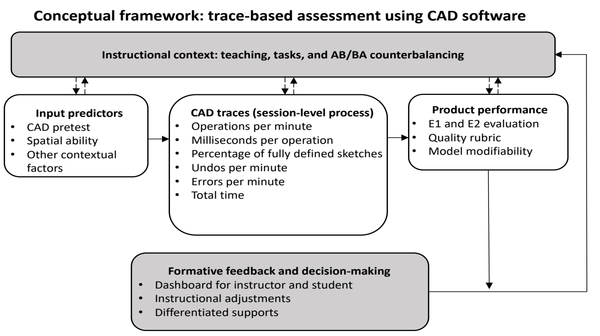


Figure 1

Conceptual framework for CAD trace-based assessment. The main flow goes from input predictors to process traces, then to product performance, and finally to formative feedback; the dashed lines indicate the sensitivity and influence of instruction on the process.

Source: Own Elaboration

Box 2

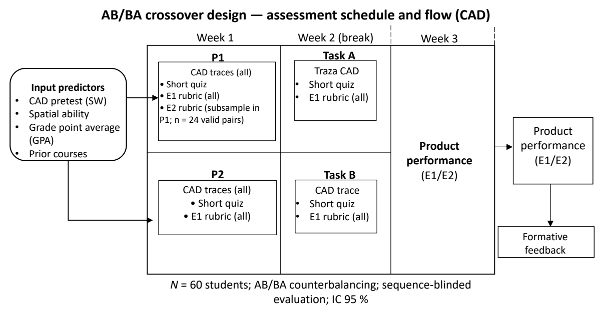


Figure 2

AB/BA cross-over design — schedule and evaluation flow (CAD)). In P1 and P2, alternating tasks are applied according to the sequence (AB or BA). In each period, CAD traces are captured, the short test is applied, and it is graded with E1 (and E2 only in a subsample of P1). N = 60 students; AB/BA counterbalancing; sequence-blinded assessment; 95% CI.

Source: Own Elaboration

Methodology

Participants and context

Sixty mechatronics students (technical degree programme at upper secondary level) participated, of whom 30 were from the morning shift (3A-MEC) and 30 from the afternoon shift (3B-MEC). In both shifts, the sequence was balanced into two groups of 15 people: AB (task A first and then task B) and BA (task B first and then A). Participation was curricular, and the data was anonymised with alphanumeric identifiers for all participants (M01–M30, V01–V30).

Unit of analysis

Although the unit of analysis consisted of 60 students (30 morning, 30 afternoon), each student completed two periods (P1 and P2). Therefore, in the analyses at the student level per period, 120 analytical observations (60 × 2) were used. It should be noted that in the Results section, n=120 refers to the observations (P1+P2) and not to the number of students.

Research design

An AB/BA crossover design was used, considering two non-contiguous periods: Period 1 (P1) in week 1 and Period 2 (P2) in week 3. Week 2 was reserved as an exposure-free interval to minimise residual effects between periods. The design allows for the estimation of task effects (A vs. B), period and sequence with intra-subject control (see Figure 2).

Allocation and blinding

Assignment to sequences AB and BA was carried out in a stratified manner by shift until 15 students per sequence were completed in each shift. The assignment list was generated before the start and was kept blind or hidden from those who graded the products, in order to avoid influencing the grading. The subsample of 30 students (50%) for the second assessor's (E2) grading was selected at random, stratified by shift (15 morning and 15 afternoon), maintaining overall parity by sequence (AB and BA), and was applied during period 1 (P1).

Tasks and materials

– **Task A** (part): parametric modelling of a part with an emphasis on creating fully defined sketches and modelling operations.

– **Task A** (part): parametric modelling of a part with an emphasis on creating fully defined sketches and modelling operations.

Environment, software, and hardware

The sessions were conducted in SolidWorks 2021 (Dassault Systèmes) on Windows 10 Pro, Dell desktop workstations, Intel Core i5-5200U CPU, 8 GB DDR3 RAM, ≥ 240 GB SSD, integrated GPU, 24" 1920×1080 monitor, and 3-button mouse.

ISSN 2531-2960
RENIECYT: 1702902

All these resources were sufficient for the planned CAD educational tasks; dedicated GPU acceleration and advanced rendering features were not used.

Obtaining and processing traces

Records or traces were obtained from the sequences generated by the software for each session, such as operations, reconstructions, undo commands, errors, and timestamps.

– **Export:** at the end of each session, the logs were exported to a CSV file per student-period, with local time zone.

– **Cleaning:** integrity (identifier or ID, chronology, duplicates) and units (millimetre, gram, second) were validated.

– **Computation:** the variables described below were summarised for each session.

– **Analytical repository:** a master file was consolidated in Excel and a syntax was prepared for SPSS v27 that allows all analyses to be reproduced.

Measures

Traces per session

The following were considered:

– Operations per minute, number of operations, number of sketches, percentage of defined sketches, number of position relationships, reconstruction time (seconds), milliseconds per operation, undo per minute, errors per minute, total time (minutes), idle time (minutes) and start of first trace (seconds).

– Interpretation: higher operations per minute and percentage of defined sketches reflect better workflow; lower milliseconds per operation, undo per minute, errors per minute, and idle time indicate smoother interaction with the tool.

Learning performance

– Rubric rating by the primary evaluator (E1, scale 0–100) applied to all analysed products.

– Rubric rating by a second evaluator (E2, scale 0–100) on a 50% subsample (30 students) in Period 1 (P1) balanced by shift (15 and 15) and with overall parity by sequence, to estimate inter-evaluator reliability.

– Short test (scale 0–100) specific to each task and period, applied at the end of each session.

.

Traceability of the analytical set for reliability and E1–E2 comparison

Of the planned subsample of 30 students with a second assessor's rating (E2) in P1, 24 cases met the analytical criteria (complete E1–E2 pairs with no missing values) and constituted the analytical set used for the estimation of reliability (CCI and α), the 'paired t-test' and the general linear model of repeated measures (within-subjects factor: measurement E1, E2; between-subjects factor: task in P1). The distribution by task in P1 was A=6 and B=18. This imbalance arose due to the exclusion from the complete list of some cases with missing data in E1 or E2; the planning maintained parity by shift and sequence, but only 24 pairs met the final analytical criteria.

Note on analytical counting

The main analysis with E1 uses 120 observations derived from 60 students measured in two periods (P1 and P2). Analyses with E2 are restricted to the subsample with complete pairs in P1 (n = 24) to emphasise that they are E1 + E2, as documented in the stated analytical traceability.

Differential attrition by sequence

The availability of E2 in P1 depended on the assigned sequence (AB/BA), with evidence of differential attrition (p = .006). Although comparability analyses showed no differences in baseline covariates between included vs. excluded cases, analyses with E2 are interpreted with caution and adjusted according to sequence.

ISSN 2531-2960
RENIECYT: 1702902

Engagement (self-report based on traces)

A brief self-report (4 items, scale 1–7) was considered

These items range from interest, attention, effort, and intention to continue, all of which were adapted and reviewed by expert judgement (three CAD teachers; 1 = strongly disagree, 7 = strongly agree).

Participation index (derived from traces)

Standardised scores (z) were averaged with positive orientation of: operations per minute, percentage of defined sketches, inverse of undo per minute, inverse of errors per minute, inverse of milliseconds per operation. Transformation [1] was applied. This transformation is applied in order to positively orient these metrics, interpreted as follows: if a lower original value is obtained, this would indicate better performance and a number of positional relationships in the assembly, with higher index values indicating greater fluidity.

Covariates

Four continuous measures were included to control for prior differences: diagnostic test of software use (SolidWorks) (0–100), spatial vision test (0–100), academic average (0–10), and prior courses in CAD/Industry 4.0 (count). All covariates were entered as continuous in their original scale in the models.

Risk label (only in P1)

An initial risk label (binary variable) was defined prior to analysis and assigned a value of 1 when any of the following conditions were met in Period 1 (P1, shown below); otherwise, a value of 0 was assigned:

– The E1 grade was in the 15th percentile or below of the P1 cohort;

– Errors per minute and downtime were both in the 85th percentile or higher (P1 cohort);

– There was a coincidence of tardiness and late delivery in the session;

– The student reported the intention to drop out; however, this rule is conservative and is therefore used as an early warning to guide specific support.

Procedure

Two sessions were considered for the evaluation, which were recorded as follows: P1 in week 1 and P2 in week 3. Week 2 was left as an interval without any exposure (no CAD activities) to reduce possible residual effects between periods. In each session, before starting, the preliminary tests were administered, and at the end, the short test and brief self-report of participation were administered, lasting approximately one minute. Product ratings were issued without the evaluators knowing the sequence assigned to the student. Each student generated two assessable products (one per period), so performance variables (e.g., E1) were also analysed at the student level for each period (120 observations). The E2 assessment was administered only in P1 to a planned subsample (30), of which 24 E1–E2 pairs (P1 subsample, n = 24) were valid.

Statistical analysis plan (SPSS v27)

1. **Mixed linear models for E1**.

A model was fitted with fixed effects for sequence (AB/BA), period (P1/P2), their interaction, and shift, with a random intercept per student. In addition, the following will be included as continuous covariates: prior testing of software use, spatial vision testing, academic average, and previous courses taken. Estimates (marginal mean differences), standard errors, p-values, and 95% confidence intervals will also be reported, with subsequent comparisons adjusted for sequence interaction by period.

2. **Inter-rater reliability**.

The two-way intraclass correlation coefficient (ICC), mixed model, consistency type (95% CI) was estimated using the E1 and E2 scores of the planned subsample of 30 students in P1. The record of a sensitivity analysis with the absolute agreement ICC was also considered.

ISSN 2531-2960
RENIECYT: 1702902

Corral-Verdugo, Alex, Sepúlveda-Romo, Adrián, Jimenez-Lopez, Eusebio and León-Rochin, German. [2025]. Development of strategic leading-edge technologies and open innovation for social transformation. Journal of Technological Development. 9 [22] 1-13: e30922113
https://doi.org/10.35429/JTD.2025.9.22.3.1.13

3. **Logistic regression for risk label (P1 only)**

The initial (binary) risk label will be the result and predictors such as: operations per minute, percentage of defined sketches, inverse of undo per minute, inverse of errors per minute, inverse of milliseconds per operation (function [1] was applied to these three metrics associated with corrections and interruptions), number of position relations, downtime, and latency to first stroke; in addition to the covariates considered (previous tests, average, and courses). Odds ratios (OR) with 95% CI, goodness of fit (including Hosmer–Lemeshow) and discriminatory performance using the area under the ROC curve (AUC) will also be reported. In addition, sensitivity and specificity will be reported at the optimal cut-off point determined by Youden's index) [2] together with the ROC curve, indicating the selected threshold and, where possible, 95% confidence intervals.

4. **Assumptions and sensitivity analysis**

– Multicollinearity. The variance inflation factor (VIF) was reviewed with a guideline threshold (< 5) and the condition number (< 30). If high collinearity is detected, centring/scaling predictors, reducing or combining variables will be considered.

– Linearity in the logit function (logarithm of the odds ratio). For continuous predictors, linearity in the logit will be verified, understood as [3] The Box–Tidwell test will be applied and, if there is evidence of non-linearity, restricted cubic splines or fractional polynomials will be considered.

– Influential observations. Standardised deviance residuals (absolute values > 3) and leverage will be inspected. Standardised deviance residuals (absolute values > 3) and leverage will be inspected. [4] stricter criterion [5] Cook's distance (strong signal > 1, initial signal [6] in addition to DFBETAs (standardised change in each coefficient when removing a case; for example [7] where p is the number of model parameters in the linear predictor (including the intercept) and n is the number of cases analysed. The cases flagged will be reviewed and any decision (to keep, transform or exclude) will be documented in the Annexes.

– Ablation by predictor groups: The regressions will be repeated, eliminating the trace variable families (flow/efficiency; corrections and interruptions; times) one by one, and compared with the complete model using the same sample. AUC (with 95% CI), sensitivity/specificity at the Youden cut-off point

$$J = \text{sensitivity} + \text{specificity} - 1 \qquad [8]$$

and changes in odds ratios (OR) (magnitude, direction and 95% CI).

5. **Missing data and outliers**

The percentage of missing data per variable and the pattern of absence will be described (where applicable, Little's test will be used). If the data lost per variable is less than or equal to 5%, a complete list analysis will be performed in its entirety; if it exceeds 5%, multiple chained imputation will be applied (at least 20 imputations and 10 iterations for each set), including all analytical variables and auxiliary variables related to the absence in the imputation model. Imputations will be made on the original scale; transformations (e.g.) and standardisations (z) will be calculated after the aforementioned imputation has been performed. The results will be combined with Rubin's rules and compared with the complete list analysis to verify the robustness of the design.

Outliers previously defined as $|z| > 3.5$ will be inspected and any decision to keep, transform or exclude them will be documented in the Annexes. In logistic models, the review of deviance residuals, leverage, Cook's distance and DFBETAs will complement this criterion.

Attrition analysis in subsample E2. Within the 30 cases planned for the second evaluator's (E2) rating in P1, those included in the analytical set (n = 24, complete E1–E2 pairs) will be compared with those excluded (n = 6) in baseline covariates (SolidWorks pretest, spatial vision test, academic average, and previous courses).

ISSN 2531-2960
RENIECYT: 1702902

Corral-Verdugo, Alex, Sepúlveda-Romo, Adrián, Jimenez-Lopez, Eusebio and León-Rochin, German. [2025]. Development of strategic leading-edge technologies and open innovation for social transformation. Journal of Technological Development. 9 [22] 1-13: e30922113
https://doi.org/10.35429/JTD.2025.9.22.3.1.13

Journal of Technological Development 9[22] 1-13: e30922113
Article

The Student's t-test for independent samples (with Levene's correction) was used, and the Mann–Whitney test was used for sensitivity. In addition, as an additional method, the distribution by task (A/B) between those included and those excluded was compared. Finally, the absence of substantive statistical differences supported the conclusion that the reduction from 30 to 24 did not introduce any systematic bias.

6. Residual effects between periods (sensitivity)

To verify the absence of residual effects between periods, two sensitivity analyses were performed, which are:

- In P2, the effect of the sequence (AB versus BA) was compared as a test for possible residual effects.
- The main analysis was repeated restricted to P1, where there is no influence from a previous period.

Ethical considerations. The study complied with institutional educational ethics regulations. Institutional informed consent was obtained, the data were anonymised and used exclusively for academic purposes.

Transparency and availability

Transparency and availability. Anonymised CSV files, a master Excel file, SPSS v27 syntax and a variable dictionary were provided, along with a step-by-step reproduction guide. The materials will be deposited in a public repository and the link/DOI will be included in the final version.

Results

Analytical set

Of the 60 students considered (P1 and P2 = 120 observations), a double assessment (E1+E2) was planned for 30 samples in P1; 24 complete pairs were analysed (as 6 were excluded due to lack of E1 or E2). In P1, the distribution presented by task was A = 6 and B = 18.

Box 3

Table 1
Subsample with E2 in P1 and grade descriptors

Group	n	E1 Media (DE)	E2 Media (DE)
Homework A	6	73.7 (4.0)	71.3 (6.0)
Homework B	18	73.9 (5.7)	74.7 (5.9)
Global	**24**	**73.9 (5.2)**	**73.8 (6.0)**

Source: Own Elaboration

Reliability and agreement among assessors (P1, n = 24)

The rating showed high consistency, with values of: α = .915; CCI = .915, 95% CI [.804, .963]; $F(23, 23) = 11.81$, $p < .001$. The mean difference (E2−E1) was −0.01 points, 95% CI [−1.31, 1.33], $t(23) = -0.02$, $p = .986$ (no systematic bias presented).

Box 4

Table 2
Reliability and paired comparison E1–E2

Metrics	Value
Cronbach's alpha	0.915
CCI (two-way, mixed, consistency)	.915 (IC 95 % [.804, .963]); F (23, 23) = 11.81; p < .001
Difference E2−E1	−0.01 (IC 95 % [−1.31, 1.33]); *t*(23) = −0.02; p = .986

Source: Own Elaboration

Viewing the agreement

The Bland–Altman test confirmed an absence of bias, with values of: bias = −0.01; 95% limits = [−6.16, 6.14]; no points outside the limits; β = 0.151, p = .222 (no proportional bias).

Box 5

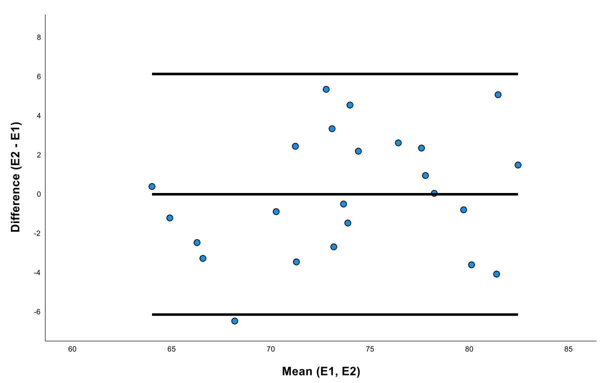


Figure 3
Bland– Altman for E1–E2 (n = 24). Average bias (E2−E1) = −0.01; 95% confidence limits = [−6.16, 6.14]; no points outside the limits; β = 0.151, p = .222.
Source: Own Elaboration

ISSN 2531-2960
RENIECYT: 1702902

Corral-Verdugo, Alex, Sepúlveda-Romo, Adrián, Jimenez-Lopez, Eusebio and León-Rochin, German. [2025]. Development of strategic leading-edge technologies and open innovation for social transformation. Journal of Technological Development. 9 [22] 1-13: e30922113
https://doi.org/10.35429/JTD.2025.9.22.3.1.13

Performance prediction

In the complete sample (P1+P2, 120 observations), a simple model with pre-tests explained approximately 61% of the variance in E1 (adjusted R^2 close to .611, F(2,117) = 94.41, p < .001). The CAD pretest was the main predictor (B = .261; 95% CI [.214, .308]; p < .001) and spatial vision contributed additionally (B = .180; 95% CI [.124, .235]; p < .001). Similarly, in the E2 subsample, the CAD pretest maintained the pattern; spatial vision was not significant.

Box 6

Table 3

Simple model for E1 (predictors: CAD pretest and spatial vision)

Results	n	Adjusted R^2	F(gl)	p (model)	Predictor	B (IC 95 %)	p
E1	120	0.611	94.41 (2,117)	< .001	Pretest de CAD	.261 [.214, .308]	< .001
					Spatial vision	.180 [.124, .235]	< .001

Source: Own Elaboration

Sequence effect in P2

No difference was detected between AB and BA when adjusting for covariates, with significance values of p = .384.

Discussion

Main findings

The present study shows that performance in CAD tasks can be anticipated based on instrumental skills, because: the SolidWorks pretest explained approximately 61% of the variance in E1 (adjusted R^2 close to .611; F(2,117) = 94.41, p < .001). Likewise, the evaluation process was robust, with high inter-rater reliability (α = .915; ICC = .849/.919) and no systematic bias between raters. A task bias (A < B) was identified, which calls for a review of existing anchors in the rubric, where no sequence effect was observed in P2.

Interpretation

The SolidWorks pretest assessment is consolidated as a good predictor of student performance because it captures the sequences of operations, procedures, and commands of the software, aligned with the quality of the product being assessed.

This supports a mechanism for close transfer to better CAD execution. The weak association between pre-tests (r = 0.16) and the lower contribution of spatial vision in E2 show that the instruments captured non-equivalent dimensions because spatial vision adds value especially in E1, while the SolidWorks (SW) pre-test offers a direct overview of the procedural domain required for modelling. Overall, the results suggest that specific preparation in the tool (SW) translates into better products, and that instructional support can be focused based on early indicators of workflow (traces) and pre-tests.

Future directions

To strengthen external validity and translate these findings into CAD curriculum improvements, the following lines of work are prioritised with explicit design and evaluation criteria:

– **Replication and sample size.** It is advisable to conduct replicas at multiple sites, calculating the sample size for an E1–E2 equivalence test, defining a target equivalence margin with standardised criteria to be established by the faculty (panel of experts).

– **Balance and design.** Balance A/B and AB/BA with stratification by shift and explicitly model the carry-over effect in the crossover design.

– **Rubric (task A).** Conduct a controlled trial, introducing anchors and examples of borderline performance, as well as evaluating the impact on bias by task and fairness by subgroups.

– **Modelling and prediction.** Employ mixed-effects models that integrate geometric complexity and reconstruction time, with cross-validation and out-of-sample evaluation to generate robust predictive performance.

ISSN 2531-2960
RENIECYT: 1702902

Corral-Verdugo, Alex, Sepúlveda-Romo, Adrián, Jimenez-Lopez, Eusebio and León-Rochin, German. [2025]. Development of strategic leading-edge technologies and open innovation for social transformation. Journal of Technological Development. 9 [22] 1-13: e30922113
https://doi.org/10.35429/JTD.2025.9.22.3.1.13

Conclusion

The study carried out in this work confirmed the validity of the design applied without sequence effects and with high inter-rater reliability. This justifies operating with a single rating, complemented by periodic audits and the application of elements to generate traceability in the study.

The SolidWorks pretest is consolidated as a key predictor of performance and indicates early detections that can be made to generate targeted support where required. Spatial vision, meanwhile, provides additional information, especially for E1, and guides timely educational planning. Furthermore, the identification of task bias (A < B) justifies reviewing and balancing the criteria of the A rubric before its next implementation.

Taken together, all these findings allow for the optimisation of assessment and teaching in CAD courses and provide a basis for future evidence-guided applications and curricular adjustments.

Appendices

Box 7

Table 4

Table A2. Paired comparison E1–E2 and effect sizes (subsample P1, n = 24)

A) Paired correlation

Par	*r*	*p*
E1 with E2	0.852	< .001

B) Paired samples t-test. (difference = E2 − E1)

Δ media	*IC 95 % de Δ*	*t*(23)	*p*
−0.01	[−1.31, 1.33]	−0.02	0.986

C) Effect sizes (paired)

Index	*Point estimate*	*IC 95 %*
Cohen's d	0.004	[−.397, .404]
Hedges' g	0.004	[−.390, .397]

Source: Own Elaboration

Box 8

Table 5

Table A3: Baseline comparability between included (valid E1–E2 pairs) and excluded (P1, planned E2 subsample) n=30

Variable (scale)	Included n=24 (Mean)	Excluded n=6 (Mean)	Test	*p*
SolidWorks pretest (0–100)	63.2	62.3	Mann–Whitney	0.67
Spatial vision (0–100)	65.8	65.4	Mann–Whitney	0.6
Overall average, GPA (0–10)	8.3	8.37	Mann–Whitney	0.49
Previous CAD/Ind. 4.0 courses (count)	1.01	1.04	Mann–Whitney	0.62
Proportion per shift (morning/evening)	—	—	$\chi^2(1) = 0.00$	1

Source: Own Elaboration

Box 9

Table 6

Table A4: Sensitivity model for E2 (P1): sequence control (AB/BA). n = 24 (complete pairs E1–E2 in P1). Dependent variable: E2_rubrica. $R^2 = .408$; adjusted $R^2 = .319$.

Source	gl	*F*	*p*	η² parcial
Corrected model	3	4.599	0.013	0.408
pretest_SW	1	10.657	0.004	0.348
pretest_vision	1	1.116	0.303	0.053
sequence_num (AB vs BA)	1	1.784	0.197	0.082
Error	20	—	—	—

Source: Own Elaboration

Statements

Conflict of Interest

The authors declare that they have no conflicts of interest.

Contribution of authors

Corral-Verdugo, Alex: Conceptualisation, Methodology, Research, Data curation, Formal analysis, Visualisation, Original draft writing, Project management, Corresponding author.

He led the study design (AB/BA crossover), coordinated trace capture and processing, performed the main statistical analyses (reliability, Bland–Altman, regressions), prepared figures/tables, and wrote the initial manuscript.

Sepúlveda-Romo, Adrián: Methodology, Software, Resources, Research, Validation, Writing–revision and editing. He configured the SolidWorks environment and the virtual classroom, supported the instrumentation for trace extraction, supervised the logistics of the sessions, validated the integrity of the data, and contributed substantive revisions to the text.

Jimenez-Lopez, Eusebio: Resources, Instrumentation, Research, Data Curation, Writing–Revision and Editing. He co-developed and refined the rubrics and short tests, supported classroom implementation and anonymisation protocols, curated the data sets, and reviewed the manuscript for clarity and pedagogical alignment.

León-Rochin, Germán: Formal analysis, validation, visualisation, writing–review and editing. Advised on statistical strategy (CCI models, mixed models, sensitivity analysis), verified reproducibility (SPSS syntax), refined figures and provided critical editorial feedback.

Availability of data and materials

The data and analysis materials (CSV/Excel, SPSS v27 syntax, and replication guide) will be openly available on Zenodo; the DOI will be provided upon publication of the article.

Funding

The research did not receive any funding.

Abbreviations

AB/BA: balanced sequence (A followed by B / B followed by A)

AUC: area under the ROC curve

BA (Bland–Altman): graph of differences and average

CAD: computer-aided design

ICC: intraclass correlation coefficient

SD: standard deviation

E1: rating of the primary assessor

E2: rating of the secondary assessor

EMMs: estimated marginal means

GLM: general linear model

95% CI: 95% confidence interval

K–S / S–W: Kolmogorov–Smirnov / Shapiro–Wilk tests

LoA: limits of agreement (Bland–Altman)

MICE: multiple imputation by chains of equations

OR: odds ratio

P1 / P2: period 1 / period 2

ROC: receiver operating characteristic
Adjusted R^2: adjusted coefficient of determination

SPSS: Statistical Package for the Social Sciences

SW: SolidWorks

VIF: variance inflation factor

References

Background

Bartlett, K. A., & Camba, J. D. (2023). Toward a broader understanding of spatial ability in CAD education. Computer-Aided Design and Applications, 24(1), 39-54. https://doi.org/10.14733/cadaps.2024.39-54

Merzdorf, H. E., Jaison, D., Weaver, M. B., Linsey, J., Hammond, T., & Douglas, K. A. (2024). Sketching assessment in engineering education: A systematic literature review. Journal of Engineering Education, 113(4), 872-893. https://doi.org/10.1002/jee.20560

Essentials

Anghel, E., Khorramdel, L., & von Davier, M. (2024). The use of process data in large-scale assessments: A literature review. Large-scale Assessments in Education, 12(1), 13. https://doi.org/10.1186/s40536-024-00202-1

Aranburu, A., Camba, J. D., Justel, D., & Contero, M. (2023). An improved explicit reference modeling methodology for parametric design. Computer-Aided Design, 161, 103541. https://doi.org/10.1016/j.cad.2023.103541

Bojčetić, N., Žeželj, D., Martinec, T., & Valjak, F. (2021). Automatized evaluation of students' CAD models. Education Sciences, 11(4), 145. https://doi.org/10.3390/educsci11040145

Company, P., Contero, M., Otey, J., & Plumed, R. (2015). Approach for developing coordinated rubrics to convey quality criteria in MCAD training. Computer-Aided Design, 63, 101–117. https://doi.org/10.1016/j.cad.2014.10.001

Contero, M., Pérez-López, D., Company, P., & Camba, J. D. (2023). A quantitative analysis of parametric CAD model complexity and its relationship to perceived modeling complexity. Advanced Engineering Informatics, 56, 101970. https://doi.org/10.1016/j.aei.2023.101970

Jaakma, K., & Kiviluoma, P. (2019). Auto-assessment tools for mechanical computer aided design education. Heliyon, 5(10). https://doi.org/10.1016/j.heliyon.2019.e02622

Mandorli, F., & Otto, H. E. (2024). Improving the learning experience within MCAD education: The provision of feedback on CAD model quality and dormant deficiency in real-time. Computer-Aided Design and Applications, 21(5), 739-758. https://doi.org/10.14733/cadaps.2024.739-758

Nobes, D. S. (2025). Automating the assessment of CAD drawings and solid models. Proceedings of the Design Society, 5, 691–700. https://doi.org/10.1017/pds.2025.10083

Otto, H. E., & Mandorli, F. (2021). Advancing formative assessment in MCAD education: The visual analytics of parametric feature-based solid models. Advanced Engineering Informatics, 48, 101308. https://doi.org/10.1016/j.aei.2021.101308

Otto, H. E., & Mandorli, F. (2025). Formative feedback improvement in MCAD education: Supporting knowledge-driven dormant deficiency detection and analysis. Computer-Aided Design and Applications, 22(4), 652–672. https://doi.org/10.14733/cadaps.2025.652-672

Paulsen, L., & Lindsay, E. (2024). Learning analytics dashboards are increasingly becoming about learning and not just analytics: A systematic review. Education and Information Technologies, 29, 14279–14308. https://doi.org/10.1007/s10639-023-12401-4

Riestra-González, M., del Puerto Paule-Ruíz, M., & Ortín, F. (2021). Early prediction of students' academic performance with learning analytics. Computers & Education, 163, 104108. https://doi.org/10.1016/j.compedu.2020.104108

Sailer, M., Ninaus, M., Huber, S. E., Bauer, E., & Greiff, S. (2024). The end is the beginning is the end: The closed-loop learning analytics framework. https://doi.org/10.1016/j.chb.2024.108305

Simonsmeier, B. A., Flaig, M., Deiglmayr, A., Schalk, L., & Schneider, M. (2022). Domain-specific prior knowledge and learning: A meta-analysis. Educational psychologist, 57(1), 31-54. https://doi.org/10.1080/00461520.2021.1939700

Susnjak, T., Ramaswami, G. S., & Mathrani, A. (2022). Learning analytics dashboard: a tool for providing actionable insights to learners. International Journal of Educational Technology in Higher Education, 19(1), 12. https://doi.org/10.1186/s41239-021-00313-7

Wang, F.-H. (2021). Interpreting log data through the lens of learning design: Second-order predictors and their relations with learning outcomes in flipped classrooms. Computers & Education, 168, 104209. https://doi.org/10.1016/j.compedu.2021.104209

Xie, C., Zhang, Z., Nourian, S., Pallant, A., & Bailey, S. (2014). On the instructional sensitivity of CAD logs. International Journal of Engineering Education, 30(4), 760–778.

Zhu Zhu, C., Leung, C. O. Y., Lagoudaki, E., Velho, M., Segura-Caballero, N., Jolles, D., ... & Klapwijk, R. (2023, April). Fostering spatial ability development in and for authentic STEM learning. In Frontiers in Education (Vol. 8, p. 1138607). Frontiers Media SA. https://doi.org/10.3389/feduc.2023.1138607

Zou, Q., Feng, H. Y., & Gao, S. (2023). Variational direct modeling: A framework towards integration of parametric modeling and direct modeling in CAD. Computer-Aided Design, 157, 103465. https://doi.org/10.1016/j.cad.2022.103465

Background

CONALEP. (2023). Mapa Curricular. Profesional Técnico Bachiller en Mecatrónica. CONALEP. https://www.conalep.edu.mx/sites/default/files/micrositios/MapaCurricular/MC_MECATRÓNICA23.pdf

Dilling, F., & Vogler, A. (2021). Fostering spatial ability through computer-aided design: A case study. Digital Experiences in Mathematics Education, 7(2), 323-336. https://doi.org/10.1007/s40751-021-00084-w

Gittinger, M., & Wiesche, D. (2024). Systematic review of spatial abilities and virtual reality: The role of interaction. Journal of Engineering Education, 113(4), 919-938. https://doi.org/10.1002/jee.20568

Secretaría de Educación Pública. (2022). *Mecatrónica*. Coordinación Sectorial de Fortalecimiento Académico. https://cosfac.sems.gob.mx/Docentes/Programas_de_estudio_del_Bachillerato_Tecnologico_Formacion_Profesional/Programas_de_estudio_de_ED_2022/DGETI/Mecatrónica.pdf

Support

Eltaief, A., Ben Amor, S., Louhichi, B., Alrasheedi, N. H., & Seibi, A. (2024). Automated Assessment Tool for 3D Computer-Aided Design Models. Applied Sciences, 14(11), 4578. https://doi.org/10.3390/app14114578