EDUVISTA
Journal of Education and Innovation
https://doi.org/10.67964/rx72zz03 Volume 1, Issue 2, July–December 2026

Applying fuzzy logic to vocational practice assessment: A critical review of methodological transfer, membership functions, and calibration

Angga Pramudika[1]*, Hartoyo[2]

[1 2] Universitas Negeri Yogyakarta, Yogyakarta, Indonesia

*Corresponding Author: anggapramudika65@gmail.com

Abstract

Vocational practice assessment requires an approach capable of representing variations in student performance that cannot always be captured through rigid grading categories. This study examines the mechanisms of *fuzzy logic* implementation, membership function applications, parameter calibration, and their suitability for vocational practical competency assessment. A *Critical Narrative Review* was conducted on 13 primary studies published between 2020 and 2026, identified through literature searches and exploratory stages. The findings indicate that fuzzy applications predominantly employ rule-based inference systems and triangular membership functions, while direct evidence from vocational practical assessment remains limited. Only two studies examined vocational education in Indonesia, and only one directly assessed industrial practice in a vocational high school. Most studies also provided insufficient information on membership boundary calibration. Limitations in the traceability of the literature search further constrained the reproducibility of the review. As its contribution, this review proposes a *hybrid crisp-fuzzy* framework that retains occupational safety requirements and mandatory competencies as strict thresholds while applying fuzzy logic to assess graded performance dimensions.

Keywords: Fuzzy logic; Vocational practice assessment; Membership functions; Parameter calibration; Psychomotor competency.

Abstrak

Penilaian praktik vokasional membutuhkan pendekatan yang mampu merepresentasikan variasi kualitas performa peserta didik yang tidak selalu dapat diukur melalui kategori nilai tegas. Penelitian ini bertujuan menganalisis mekanisme penerapan *fuzzy logic*, penggunaan fungsi keanggotaan, kalibrasi parameter, serta kesesuaiannya untuk penilaian kompetensi praktik vokasional. Kajian ini menggunakan pendekatan *Critical Narrative Review* terhadap 13 studi primer yang diterbitkan pada 2020–2026, yang diidentifikasi melalui penelusuran literatur dan tahap eksploratif. Hasil menunjukkan bahwa penerapan fuzzy didominasi sistem inferensi berbasis aturan dan fungsi keanggotaan triangular, sedangkan bukti langsung dalam penilaian praktik vokasional masih terbatas. Hanya dua studi membahas pendidikan vokasional di Indonesia dan satu secara langsung menilai praktik industri SMK. Sebagian besar studi juga belum melaporkan kalibrasi batas fungsi keanggotaan secara memadai. Keterbatasan ketertelusuran penelusuran literatur turut membatasi reproduksibilitas kajian. Sebagai kontribusi, kajian ini mengusulkan kerangka *hybrid crisp-fuzzy* yang mempertahankan persyaratan keselamatan dan kompetensi wajib sebagai batas tegas, sementara fuzzy digunakan untuk menilai kualitas performa bertingkat.

Kata Kunci: *Fuzzy logic*; Penilaian praktik vokasional; Fungsi keanggotaan; Kalibrasi parameter; Kompetensi psikomotor.

© 2026 by Author(s). Submitted for possible open access publication under the terms and conditions of the Attribution 4.0 International (CC BY 4.0) license (https://creativecommons.org/licenses/by/4.0/).

Introduction

Assessment in vocational education needs to reflect the quality of students' performance when carrying out tasks, rather than merely their mastery of theoretical knowledge. Practical performance is shaped by several interrelated indicators, including procedural accuracy, output quality, speed, consistency, and compliance with occupational safety requirements. These indicators do not necessarily move in the same direction; therefore, simply summing scores or applying a single categorical threshold risks oversimplifying competency quality. Namlı and Şenkal (2018) examined the use of a fuzzy system to assess the programming performance of 336 university students using a Mamdani model that processed two inputs into a single output. The fuzzy-based assessment results differed significantly from those of the absolute evaluation used as a comparison. This finding indicates that assessment decisions may change when relationships among performance components are considered through fuzzy rules. In workshop practice, an error at one stage may even compromise safety and the quality of the final product.

Another problem arises when students' scores fall close to the passing threshold. Ivanova and Zlatanov (2019) demonstrated that, on a test with a maximum score of 60, a student scoring 29 was classified as failing, whereas scores of 30 or 31 could result in a passing status, despite the negligible difference in achievement. They applied fuzzy functions to address such borderline cases and refine grade categorization without statistically altering the group mean. This issue is relevant to vocational education because practical assessment also involves categorical decisions about performance that varies by degree. A student may demonstrate an almost correct procedure while still differing in precision, neatness, or work efficiency. Assessment should therefore consider the degree of achievement for each indicator rather than relying solely on rigid passing thresholds. This model allows for more nuanced assessment than a simple pass-or-fail decision.

The use of fuzzy logic in educational assessment has developed with a range of objectives and mechanisms. Amelia et al. (2019) reviewed 38 articles published between 2008 and 2018 and mapped the use of fuzzy logic in assessing student performance across various educational contexts. At the process-analysis level, Cui et al. (2020) applied an *Adaptive Neuro-Fuzzy Inference System* (ANFIS) to technology-based assessment log data to extract characteristics of problem-solving processes and refine fuzzy rules. Meanwhile, Wu et al. (2022) combined the *Analytic Hierarchy Process* and *Fuzzy Comprehensive Evaluation* to evaluate university students' learning by weighting indicators and conducting graded assessments. These findings demonstrate that fuzzy logic does more than convert scores into categories; it can also integrate multiple indicators and represent learning processes. These differences in function

indicate that model selection should be guided by the characteristics of the assessment object and the intended assessment objectives.

The application of fuzzy logic has also begun to receive attention in the contexts of skills assessment and vocational education, although the objects being assessed are not necessarily equivalent to occupational practical competencies. Tang et al. (2021) combined fuzzy logic with Gagné's learning hierarchy to assess mathematical skills. Zhang and Wang (2024) developed a fuzzy approach to validate changes in performance in session-based vocational training, while Wei (2024) used a fuzzy evaluation model to assess mathematics learning in higher vocational education. Although these studies extend the contexts in which fuzzy logic is applied, they do not provide a focused critical synthesis of how fuzzy mechanisms, membership function forms, and parameter calibration are used to assess practical skills. The review by Amelia et al. (2019) also focuses broadly on student performance assessment rather than on the methodological transfer of fuzzy approaches to vocational practical assessment. These contextual differences require their findings to be interpreted cautiously before being applied to workshop-based assessment.

Accordingly, the issue that needs to be addressed is no longer simply whether fuzzy logic can be applied to educational assessment, but rather the extent to which the selected mechanisms and membership functions are appropriate for assessing vocational practical competencies. This review aims to examine patterns of fuzzy logic application, compare the mechanisms and membership functions employed, and evaluate the reported bases for calibration, rule transparency, and validation. This focus is important because practical assessment decisions concern observable skills, the quality of work outcomes, and safety requirements that cannot always be compensated for by other indicators. Through a critical examination of the available evidence, this review is expected to establish the limits of transferring findings from academic assessment to practical assessment and identify the methodological requirements for developing more defensible vocational assessment practices, while also clarifying directions for future research.

Method

This study adopts a *Critical Narrative Review* approach to examine the application of *fuzzy logic* in vocational practical assessment, the types of membership functions employed, the methods used to determine their parameters, and the strengths and limitations of this approach. Rather than merely summarizing research findings, the review compares methodological choices and evaluates their relevance to performance assessment. The literature was searched through Scopus, Google Scholar, Garuda, and ERIC using English- and Indonesian-language terms related to *fuzzy logic*, *fuzzification*, *membership functions*, vocational education, practical assessment, and skills assessment. The search terms were combined using queries such as "fuzzy logic" AND "vocational

assessment", "fuzzy inference" AND "vocational education", "logika fuzzy" AND "penilaian praktik", and "fuzzifikasi" AND "penilaian keterampilan". Publications were restricted to the period 2017–2026. The choice of approach and search scope was guided by the review questions and objectives (Grant & Booth, 2009; Snyder, 2019).

This review addresses three research questions: (1) How is fuzzification applied in vocational practice assessment? (2) Which membership functions are most widely used, and how are their parameters determined? (3) What are the strengths and limitations of fuzzy-based assessment compared with conventional assessment in the vocational context? Table 1 presents the search strings used in the documented Phase 3 searches.

Table 1. Search strings used in the documented Phase 3 searches (3 November 2026)

No	Database	Search string	Access date	Recorded results
1	Google Scholar	"logika fuzzy" "penilaian praktik"	3 November 2026	0
2	Google Scholar	"fuzzifikasi" "penilaian praktik"	3 November 2026	0
3	Google Scholar	"logika fuzzy" "SMK"	3 November 2026	More than 100; first 100 logged
4	Google Scholar	"fuzzy logic" "practical assessment"	3 November 2026	More than 100; first 100 logged
5	Google Scholar	"fuzzy logic" "practicum assessment"	3 November 2026	4
6	Google Scholar	"fuzzy inference" "vocational assessment"	3 November 2026	4
7	Semantic Scholar	fuzzy logic + vocational assessment	3 November 2026	About 8,260; first 100 logged
8	Semantic Scholar	fuzzy logic + practicum assessment	3 November 2026	About 8,020; first 100 slots logged (91 unique because of repetition on page 6)
9	ERIC	fuzzy logic + vocational assessment	3 November 2026	53 (complete)

Source: Author's synthesis

Studies were selected if they were journal articles or conference proceedings published between 2017 and 2026, written in Indonesian or English, and addressed *fuzzy logic* in educational assessment, skills assessment, or vocational practice. Blogs, non-peer-reviewed writings, undergraduate theses, master's theses, studies outside the research focus, and publications for which the full text could not be obtained were excluded. The selection process involved identification, duplicate removal, title and abstract screening, full-text retrieval, and eligibility assessment. The selection process is

presented in Figure 1, adapting the structure of the PRISMA 2020 flow diagram (Page et al., 2021). Figure 1 adapts the structure of the PRISMA 2020 flow diagram for presentation only; this review is a critical narrative review and does not claim compliance with PRISMA 2020.

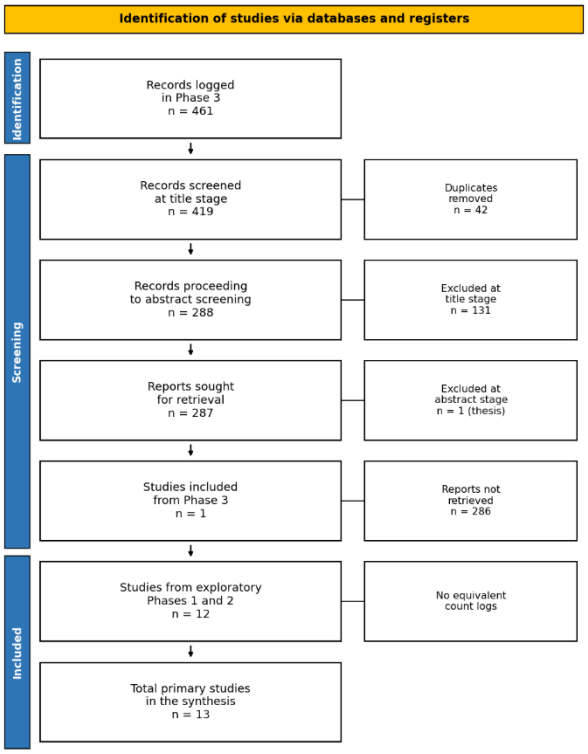

Figure 1. Study selection flow diagram based on PRISMA 2020

Source: Adapted from Page et al. (2021).

As shown in Figure 1, Phase 3 selection proceeded in stages from 461 records logged: 42 duplicates were removed, 419 were screened at the title stage, and 131 were excluded (16 outside 2017 to 2026, 6 on general learning outcomes, and 109 on different objects or contexts), leaving 288 to proceed to the abstract stage. At the abstract stage, 1 record was excluded because it was a thesis; 287 reports were sought for full-text retrieval, but 286 could not be obtained, and only 1 study was included from Phase 3 (Pahtoni and Arifin, 2023). The other 12 primary studies were identified in the exploratory Phases 1 and 2, for which equivalent count logs are not available, so the selection process cannot be reproduced from a single documented search.

A tiered rubric appraises the 13 primary studies on four dimensions: boundary calibration, rule or weight transparency, sample adequacy, and external validation. Tiers: High means complete and checkable; Medium means partial or reliant on standards and expert consensus; Low means minimal or only an illustrative example. Sample benchmarks: High means hundreds to thousands, Medium means 70 to 100 with additional testing, Low means fewer than 70 from one institution or no empirical sample; for experts, Medium means 10 to 20, Low means fewer than 10. External

validation here means comparing fuzzy results with a benchmark independent of the system itself, such as manual assessment, a standardized test score, or data from a different session or cohort; internal system performance tests, such as model accuracy on its own data or functionality testing, do not count as external validation. Not Reported means the source did not report that information. This rubric was compiled for this review, not the Critical Appraisal Skills Programme (CASP) or Joanna Briggs Institute (JBI) tools; the two comparator reviews and the general reference were not appraised. Izzati et al. (2025) were verified through the Oulu copy: triangular, 30 students, MAPE 1.35 percent. The studies were then synthesized narratively by comparing their assessment objectives, fuzzy mechanisms, membership function types, parameter-setting procedures, and reported limitations. The analysis also distinguished direct evidence from vocational education from evidence transferred from general educational contexts. This critical synthesis aimed to identify methodological patterns and evaluate their suitability for assessing practical competencies.

Results and Discussion

Mapping the research landscape of fuzzy logic in educational assessment

The main corpus of this review comprises 13 studies published between 2020 and 2026. The publication distribution varies across years: one article was published in 2020, three in 2022, two in 2023, four in 2024, one in 2025, and two in 2026. No articles were published in 2021. The highest concentration of publications occurred in 2024, covering research on online tests, industrial practice, vocational education performance, and learning quality. This pattern indicates that the application of *fuzzy assessment* has expanded to address various educational issues, although the number of studies in the corpus does not demonstrate consistent growth across years. This distribution provides an initial overview of developments in the research examined and serves as a basis for understanding the variations in contexts, assessment objects, and approaches employed in applying fuzzy logic to educational assessment.

Geographically, the 13 studies cover nine countries: Indonesia (four studies), China (two studies), and Albania, Taiwan, Malaysia, Italy, Türkiye, Saudi Arabia, and Bulgaria (one study each). The Indonesian studies address online quiz integrity, industrial practice assessment in vocational high schools, prediction of university students' academic achievement, and the evaluation of learning outcomes among vocational high school students. Studies from China focus on vocational education performance and learning quality among students in higher vocational education. Meanwhile, studies from other countries cover biomedical assessment, *joinery* skills, mathematics, group projects, distance learning, academic prediction, and language tests. This distribution indicates that the use of fuzzy logic in educational assessment has extended across diverse international contexts. Details of the countries, research

objects, data characteristics, approaches, and main findings of all articles are presented in Table 2.

Table 2. Characteristics, countries, and main findings of 13 primary studies on fuzzy assessment (2020–2026)

No.	Study	Country	Research object	Data and approach	Main findings
1	Maraj et al. (2026)	Albania	Assessment of university students' academic performance in biomedical courses.	50 students; four fields: *biomechanics, anatomy, physiology,* and *kinesiology*; Mamdani FIS with 10 rules.	The correlation between fuzzy results and conventional averages reached r = 0.91, with an MAE of 0.34 and an RMSE of 0.42. In one example, a conventional average score of 7.00 yielded a fuzzy score of 6.42.
2	Nirawana et al. (2026)	Indonesia	Assessment of student performance and examination integrity in a *web-based smart quiz* system.	58 students and 81 quiz attempts; four inputs: score percentage, examination integrity, number of attempts, and time ratio; Mamdani FIS.	The 81 attempts were classified as 33 *Superior*, 39 *Good*, 7 *Enough*, and 2 *Needs Coaching*. The risk analysis categorized 79 attempts as *Safe*, while 2 required further attention.
3	Tsai et al. (2020)	Taiwan	Development of *joinery* skills training indicators for the regional Taiwan National Skills Competition.	14 experts for Fuzzy Delphi and 15 experts for FAHP; six main indicators and 27 sub-indicators.	*Measurements* received the highest weight at 19.6%, followed by *finishing and appearance* at 18.9%. All six main indicators were retained.
4	Aziz et al. (2022)	Malaysia	Ranking of university students' group projects.	Three projects; four main criteria; 16 subcriteria; three experts; Fuzzy Analytic Hierarchy Process (FAHP).	The third project alternative received the highest weight of 0.5379 and a manual rubric score of 83.33%. The FAHP ranking was consistent with the results of the manual rubric assessment.
5	Doz et al. (2022)	Italy	Assessment of mathematical knowledge among upper secondary	2,279 students; written examination scores, oral examination scores, and INVALSI results; two fuzzy	The hypothetical mean scores were 5.73 for COG and 5.56 for MOM, lower than the written examination

			school students in grade 13.	approaches: *Center of Gravity* (COG) and *Mean of Maximum* (MOM).	score of 6.50 and the oral examination score of 6.59. The correlations between written and oral examination scores and INVALSI were ρ = 0.384 and ρ = 0.380, respectively.
6	Özseven and Çağman (2022)	Türkiye	Development of a student performance evaluation model for distance learning.	The model incorporated behavior throughout one semester, the weights of multiple-choice distractors, and characteristics of incorrect answers; the article presented an application example involving three students.	The model integrated semester-long activity and the characteristics of incorrect answers into performance evaluation. The article placed greater emphasis on model development and application examples than on large-scale empirical testing.
7	Hegazi et al. (2023)	Saudi Arabia	Reasoning and prediction of students' academic performance based on attendance and final examination scores.	40 students from two classes, each comprising 20 students; one class was used for training and the other for testing; *Fuzzy Propositional Model* (FPM).	The FPM model produced an MSR of 0.062242, lower than the linear regression result of 0.076. This indicates better predictive performance on the dataset used in the study.
8	Glushkova et al. (2024)	Bulgaria	Evaluation of university students' foreign language proficiency through online tests.	78 students in the final examination and 36 students in a 70-item multiple-choice test; bell-shaped fuzzy functions.	Fuzzification was used to adjust scores around category boundaries. The evaluation compared fuzzy results with classical assessment and examined changes in the group score distribution.
9	Pahtoni and Arifin (2023)	Indonesia	Monitoring and assessment system for industrial practice among	24 supervising teachers and students; five inputs: attendance, neatness, attitude, skills, and knowledge;	System testing yielded a functional suitability rate of 100% and usability of 84.8%. These figures indicate the system's functional

			vocational high school students.	Tsukamoto FIS with 15 rules.	and usability suitability.
10	Zhang and Wang (2024)	China	Evaluation of vocational education performance and training effectiveness.	Performance data from students and instructors; 60 students; *Fuzzy-based Performance Improvement Validation Method* using fuzzification over continuous intervals and comparisons across sessions.	The model compared changes in student performance and instructor skills across sessions, as well as satisfaction factors, to identify performance changes and training adjustment needs.
11	Izzati et al. (2025)	Indonesia	Prediction of academic achievement among higher education students.	30 students; five inputs: attendance, assignments, midterm examination, final examination, and attitude; Mamdani FIS with 243 rules.	The system classified performance into *poor*, *fair*, and *good* categories. Testing produced a *Mean Absolute Percentage Error* (MAPE) of 1.35%.
12	Abiyyu and Widayanti (2023)	Indonesia	Evaluation of learning outcomes among students at SMK PGRI 3 Sidoarjo.	20 students; five criteria: social attitude, spiritual attitude, skills, knowledge, and absenteeism; Fuzzy AHP.	Four ranking results differed from those generated by the previous system. The reported level of agreement was 80%; the new system used five criteria, whereas the previous system used two.
13	Wu et al. (2024)	China	Development of a learning quality evaluation model for students in higher vocational education.	Fuzzy Comprehensive Evaluation (FCE) combined with AHP; evaluation indicators were arranged hierarchically and involved multiple assessors.	The model integrated self-assessment, peer assessment, core group member assessment, and lecturer assessment. The six dimensions covered theoretical knowledge, operational skills, ideological and political quality, social quality, learning attitude, and learning methods.

Source: Author's synthesis based on the 13 primary studies included in the review.

The assessment objects in the corpus demonstrate broad coverage. Maraj et al. (2026) assessed student performance in biomedical courses, Doz et al. (2022) processed mathematics scores among upper secondary school students, while Glushkova et al. (2024) evaluated foreign language proficiency through online tests. Other studies examined examination integrity and quiz performance (Nirawana et al., 2026), group projects (Aziz et al., 2022), learning behavior in distance education (Özseven & Çağman, 2022), and the prediction of academic achievement (Hegazi et al., 2023; Izzati et al., 2025). In vocational education, Zhang and Wang (2024) examined training performance, while Wu et al. (2024) developed a learning quality evaluation framework. Two other studies addressed *joinery* skills and industrial practice in vocational high schools. This diversity indicates that *fuzzy assessment* has been used to evaluate various forms of educational achievement, ranging from academic knowledge to vocational skills and learning quality.

Based on their primary functions, the 13 studies can be classified into three patterns of application. Five studies converted raw scores into graded categories, five combined multiple indicators into a single evaluation outcome, and three used fuzzy logic for reasoning or prediction. The first pattern appeared in biomedical assessment, online quizzes, mathematics, language tests, and vocational performance assessment. The second was evident in the development of a *joinery* skills index, project ranking, vocational high school learning outcome evaluation, semester-long behavioral assessment, and learning quality evaluation. The third pattern comprised academic prediction using the *Fuzzy Propositional Model*, the Tsukamoto system for industrial practice, and student achievement prediction using Mamdani. This classification demonstrates that *fuzzy logic* serves different assessment purposes. Consequently, findings from one type of application cannot automatically be assumed to apply to other forms of assessment.

The extent to which the studies relate to vocational practical skills also varies. Tsai et al. (2020) developed *joinery* skills indicators for the Taiwan skills competition through six indicators and 27 sub-indicators. Pahtoni and Arifin (2023) developed a monitoring and assessment system for industrial practice among vocational high school students based on attendance, neatness, attitude, skills, and knowledge. These two studies provide the most direct evidence in the corpus concerning procedural or psychomotor performance. The other 11 studies primarily focus on cognitive and academic achievement, projects, examination integrity, or learning quality and are therefore treated as transfer evidence. Although Abiyyu and Widayanti (2023) examined vocational high school students in Indonesia, their study evaluated general learning outcomes rather than industrial practice competencies. Thus, findings from academic and administrative studies provide a methodological foundation but cannot replace empirical evidence on practical skills assessment.

The characteristics of the data and research outputs also vary considerably. Doz et al. (2022) analyzed 2,279 students, whereas Hegazi et al. (2023) used two classes, each comprising 20 students. Other studies involved 20 vocational high school students, 24 supervising teachers and students, 30 university students, 50 biomedical students, 58 students with 81 quiz attempts, and 60 vocational education students. Tsai et al. (2020) involved expert panels in two stages, whereas Aziz et al. (2022) assessed three projects using input from three experts. The research outcomes also involved different measures, including correlation coefficients, prediction errors, indicator weights, ranking agreement, and system suitability. This variation in units of analysis and research outputs indicates that the success of *fuzzy assessment* depends on its intended purpose and application context. Therefore, the figures reported in individual studies need to be interpreted in relation to the measurement functions employed rather than compared as if they represented a single, uniform measure of effectiveness.

Mechanisms and patterns of fuzzy logic application in educational assessment

The application of *fuzzy logic* in educational assessment involves various mechanisms depending on the objectives and characteristics of the assessment objects. Theoretically, the concept of *fuzzy sets*, introduced by Zadeh (1965), provides the foundation for representing the degree to which an object belongs to categories with no strictly defined boundaries. One application is the *Fuzzy Inference System* (FIS), with the Mamdani model presented by Mamdani and Assilian (1975) and the Takagi-Sugeno model (Takagi & Sugeno, 1985), which processes inputs through membership functions and IF–THEN rules to generate outputs. Based on these mechanisms, the literature analyzed in this review can be classified into three approaches: *rule-based FIS*; direct fuzzification and *Fuzzy Comprehensive Evaluation* (FCE); and *Fuzzy Multi-Criteria Decision Making* (MCDM). Direct fuzzification converts numerical scores into membership degrees without an inference engine, FCE aggregates membership degrees across several evaluation dimensions, while MCDM uses fuzzy numbers to determine criterion weights and priorities. The differences in mechanisms, indicators, and outputs are summarized in Table 3 as a basis for comparing the functions of each approach and assessing the limitations of their application to vocational practical assessment.

Table 3. Fuzzy assessment mechanisms and configurations in educational literature

Study	Mechanism and configuration	Assessment indicators or objects	Data or sample	Reported processes and findings
Maraj et al. (2026)	Mamdani FIS; 4 inputs; 10 rules; trapezoidal functions	Biomechanics, anatomy, physiology, and kinesiology	50 students	A uniform score of 7 produced a fuzzy score of 6.42, compared with a

				conventional average of 7.00.
Nirawana et al. (2026)	Mamdani FIS; 4 inputs; triangular and trapezoidal functions	Score percentage, integrity, number of attempts, and time ratio	58 participants; 81 quiz attempts	Generated the categories *Superior*, *Good*, *Enough*, and *Needs Coaching*; 79 attempts were classified as safe, while 2 required attention.
Doz et al. (2022)	Fuzzy inference; triangular functions	Written and oral mathematics examination scores	2,279 participants	The *Center of Gravity* average was 5.73 and the *Mean of Maximum* average was 5.56; correlations with INVALSI scores were $\rho = 0.384$ and $\rho = 0.380$, respectively.
Özseven and Çağman (2022)	FIS; 7 linguistic categories; triangular functions	One-semester behavior, multiple-choice distractor weights, and incorrect answers	Application example involving 3 students	Processed behavioral and test-performance information through seven linguistic expressions.
Hegazi et al. (2023)	*Fuzzy Propositional Model* (FPM); triangular functions	Final examination scores and attendance rates	40 participants; 2 classes	Produced a *mean squared residual* of 0.062242, lower than the linear regression result of 0.076.
Pahtoni and Arifin (2023)	Tsukamoto FIS; 5 inputs; 15 rules; monotonic functions	Attendance, neatness, attitude, skills, and industrial practice knowledge	24 respondents	Generated a web-based practical assessment system; functional testing reached 100% and usability was 84.8%.
Izzati et al. (2025)	Mamdani FIS; 5 inputs; 243 rules; triangular functions; centroid defuzzification	Attendance, assignments, midterm examination, final examination, and attitude	30 students	Generated predictions of academic achievement with a MAPE of 1.35%.
Glushkova et al. (2024)	Direct fuzzification; bell-shaped functions	Test scores and five achievement levels	78 participants in the final examination; 36	Converted examination scores into membership

			participants in the context of a 70-item multiple-choice test	degrees and compared fuzzy score distributions with conventional scores.
Zhang and Wang (2024)	Fuzzification over continuous score intervals	Student and instructor performance scores across several sessions	60 participants	Used a fuzzy approach to evaluate improvements in training performance based on scores across sessions.
Wu et al. (2024)	FCE and AHP; membership matrix; three-level hierarchy	Six dimensions of the vocational learning quality index	Not reported	Generated an evaluation vector based on criterion weights and membership degrees.
Tsai et al. (2020)	Fuzzy Delphi and FAHP; triangular and trapezoidal fuzzy numbers	Six indicators and 27 sub-indicators of *joinery* skills	14 experts for Fuzzy Delphi; 15 experts for FAHP	Determined the indicator structure and weights; measurement received a weight of 19.6%, while finishing and appearance received 18.9%.
Aziz et al. (2022)	FAHP; four criteria and 16 subcriteria	Group project criteria and subcriteria	3 experts	Generated priority weights; the third alternative received the highest weight of 0.5379.
Abiyyu and Widayanti (2023)	Fuzzy AHP; five criteria	Social attitude, spiritual attitude, skills, knowledge, and absenteeism	20 students	Generated an evaluation based on five criteria, with 80% agreement with the previous system.

Source: Compiled from the literature.

Table 3 shows that *rule-based FIS* is the approach with the greatest diversity in input and rule configurations. Maraj et al. (2026) employed a Mamdani model with four inputs—biomechanics, anatomy, physiology, and kinesiology—processed through ten rules. The results showed that participants with a uniform score of 7 received a fuzzy score of 6.42, lower than the conventional average of 7.00. Nirawana et al. (2026)

combined score percentages, integrity, number of attempts, and time ratios to evaluate online quiz performance. Meanwhile, Pahtoni and Arifin (2023) applied the Tsukamoto model with five indicators and 15 rules to assess industrial practice. These comparisons demonstrate that FIS can integrate different indicators into a single decision. However, this flexibility does not automatically guarantee assessment accuracy, as the results depend on the selection of variables, membership boundaries, and rules designed by the developers.

Differences in FIS configurations are also evident in their processing objectives and output types. Doz et al. (2022) used fuzzy inference to process written and oral mathematics examination scores from 2,279 participants. The *Center of Gravity* average was 5.73 and the *Mean of Maximum* average was 5.56, whereas the written and oral examination scores were 6.50 and 6.59, respectively. Both fuzzy results correlated with INVALSI scores at 0.384 and 0.380, respectively. Özseven and Çağman (2022) incorporated one-semester behavior, multiple-choice distractor weights, and incorrect answers through seven linguistic categories. Hegazi et al. (2023) used a *Fuzzy Propositional Model* to predict performance based on final examination scores and attendance. The model produced a *mean squared residual* of 0.062242, lower than the linear regression result of 0.076. These findings indicate that FIS can be used for categorization and prediction. However, successful academic prediction does not establish the accuracy of occupational skills measurement.

The second approach comprises direct fuzzification and FCE, both of which use membership degrees but serve different processing functions. Glushkova et al. (2024) used bell-shaped membership functions to convert examination scores into five achievement categories. The study compared fuzzy and conventional score distributions, demonstrating how scores near category boundaries can be represented through membership degrees. Zhang and Wang (2024) applied fuzzification to continuous score intervals to evaluate changes in participant performance across several training sessions. In contrast, Wu et al. (2024) combined FCE with AHP through a three-level index hierarchy and six dimensions of learning quality evaluation in vocational education. This distinction is important: direct fuzzification transforms the representation of scores, whereas FCE aggregates multiple evaluation dimensions. Because these three studies differ in their assessment objects and outputs, their findings cannot be directly treated as evidence of the validity of individual practical competency assessment.

The third approach, *Fuzzy MCDM*, focuses more on weighting and ranking criteria than on directly measuring student performance. Tsai et al. (2020) combined *Fuzzy Delphi* and the *Fuzzy Analytic Hierarchy Process* to establish six indicators and 27 sub-indicators of *joinery skills*. The results showed that the measurement indicator received the highest weight of 19.6%, followed by finishing and appearance at 18.9%. The

study's contribution lies in developing an indicator structure and assigning weights based on expert judgments, rather than creating an automated system for grading student performance. Aziz et al. (2022) used FAHP to determine the priorities of four criteria and 16 subcriteria for group projects. Abiyyu and Widayanti (2023) applied the method to five criteria for evaluating vocational high school students' learning outcomes and reported 80% agreement with the previous system. Thus, MCDM helps establish assessment priorities, but criterion weights should not be equated with evidence of accurate practical competency measurement.

When examined in terms of their ultimate functions, applications of *fuzzy logic* in the literature can be mapped into three patterns: converting scores into categories, integrating multiple indicators, and reasoning or predicting performance. The first pattern translates numerical scores into linguistic categories such as low, medium, and high. The second integrates several assessment dimensions with different weights or levels of importance. The third uses relationships among variables to estimate student outcomes or levels of achievement. Through a systematic review of 109 studies and a meta-analysis of 46 studies, Tuan et al. (2025) demonstrated the wide range of applications of fuzzy logic in student performance assessment. Thuhra and Kaur (2026), in their review of 26 studies published between 2010 and 2025, identified the predominance of Mamdani models, triangular and trapezoidal functions, and centroid defuzzification. These two reviews reinforce the mapping presented in Table 3. However, the predominance of particular configurations does not automatically indicate that they are the most appropriate methods for every type of competency.

The most important implication of the mapping in Table 3 is that the success of *fuzzy logic* in academic assessment is insufficient to establish its suitability for vocational practical assessment. Evidence directly related to procedural or psychomotor skills remains limited. Tsai et al. (2020) developed indicators and weights for *joinery* skills, while Pahtoni and Arifin (2023) developed an industrial practice assessment system in vocational high schools using five indicators. As historical evidence outside the review's analytical period, Ghozali and Haryanto (2016) also developed a fuzzy-based performance assessment instrument for vocational high school students' practical skills in programmable control systems. These three studies make different contributions: indicator development, assessment system development, and practical assessment instrument development. However, these contributions are not sufficient to demonstrate the superiority of fuzzy logic in measuring actual competencies. Its application requires observable indicators, parameters calibrated against occupational standards, and safety requirements that cannot be compensated for by scores on other indicators.

Membership functions, calibration, limitations, and implications of fuzzy logic in vocational practical assessment

Membership functions determine how scores or assessments are translated into linguistic categories; therefore, their selection is not merely a technical matter. Across the 13 primary studies, triangular functions were reported in eight studies and trapezoidal functions in three, while two studies used more than one function type. However, these figures need to be interpreted cautiously because not all of them represent membership functions applied to student scores. In the three *multi-criteria decision-making* (MCDM) studies, for example, *triangular fuzzy numbers* were used to accommodate expert judgments in criterion weighting. This application differs from membership functions that map performance scores to specific categories. Therefore, the predominance of triangular functions in the corpus cannot be directly interpreted as evidence that this form is the most appropriate for assessing vocational practical competencies.

Differences in the purposes of membership functions become particularly important when research findings are applied to practical assessment. Tsai et al. (2020) used *Fuzzy Delphi* and the *Fuzzy Analytic Hierarchy Process* to develop assessment indicators and weights for *joinery* skills. Meanwhile, Pahtoni and Arifin (2023) applied the monotonic Tsukamoto functions to process five indicators of industrial practice among vocational high school students. Both studies are more closely related to procedural skills assessment than studies evaluating examination results or academic achievement. Suswanto et al. (2025) further demonstrated the relevance of fuzzy approaches in technical education through a longitudinal fuzzy model for evaluating the effectiveness of practical learning. However, evaluating learning effectiveness is not equivalent to measuring individual psychomotor competencies. Accordingly, methodological transfer must consider the object being assessed rather than merely the shared use of fuzzy logic.

The accuracy of membership functions is also determined by category boundaries and the widths of overlaps between categories. Among the 13 primary studies, only Izzati et al. (2025) reported complete membership function ranges: low at 0–60, medium at 50–80, and high at 70–100. These ranges produce ten-point overlaps between the low–medium and medium–high categories. Adjustments to the attendance and attitude indicators resulted in different overlap widths. By contrast, the score ranges of A (91–100) and B (83–90) reported by Abiyyu and Widayanti (2023) represent crisp grade categories rather than fuzzy membership function ranges. This distinction indicates that reporting linguistic labels alone is insufficient. Researchers need to present the parameters of each function so that membership degrees, category transition points, and their assessment implications can be examined transparently.

Parameter determination across the corpus reveals varying approaches that have not yet been accompanied by sufficiently robust calibration standards. Eight studies explained the basis for parameter selection, ranging from institutional standards and

assessment principles to expert consensus, variable score ranges, and AHP or FAHP weighting. The remaining five studies did not adequately explain their parameter-setting procedures. Through a review of membership function applications in the electrical engineering field, Zangeneh et al. (2020) showed that parameter selection depends heavily on system characteristics and designer knowledge; no single universal procedure applies to all systems. This technical finding is relevant to interpreting the educational corpus: expert judgment can serve as a starting point, but it does not automatically guarantee objectivity. Therefore, parameters need to be tested against empirical data, performance standards, and inter-rater agreement before being used as the basis for competency decisions. A summary of membership function patterns and calibration practices is presented in Table 4.

Table 4. Synthesis of membership functions and calibration in the 13 primary studies

Aspect	**Findings in the corpus**	**Methodological implications**
Function types	Triangular functions were reported in 8 studies, while trapezoidal functions were reported in 3 studies. Two studies used more than one function type.	Frequency of use does not automatically indicate which function is most appropriate for practical assessment.
Use in MCDM	Three studies used *triangular fuzzy numbers* for weighting based on expert judgments.	Fuzzy weights must be distinguished from membership functions that map student scores to categories.
Vocational practical studies	One study developed indicators and weights for *joinery* skills; one applied Tsukamoto to industrial practice in vocational high schools.	Direct evidence on membership functions in practical assessment remains limited.
Reporting of ranges	Only one study reported complete membership function ranges.	Ranges, transition points, and category overlaps need to be explicitly reported.
Basis for parameter determination	Eight studies reported the basis for parameter determination; five did not.	Expert consensus and institutional standards need to be followed by empirical testing.
Calibration	Nine studies received a low rating for boundary calibration.	Parameters need to be tested against performance benchmarks and inter-rater consistency.
Validation	Evidence of external validation remains limited; internal testing alone is insufficient to establish assessment validity.	Systems need to be compared with analytical rubrics or independent competency standards.

Source: Author's analysis.

Calibration involves more than selecting triangular or trapezoidal functions; it requires testing whether the functions adequately represent the characteristics of the

data being assessed. Bhattacharyya and Mukherjee (2020) proposed approaches to evaluating membership functions through nonlinear regression and algorithmic methods, comparing them across several cases. Their contribution demonstrates that membership functions can be evaluated and adjusted rather than accepted solely on the basis of designers' intuition. Adánez et al. (2019) also showed that multidimensional membership functions optimized using genetic algorithms can improve nonlinear system identification compared with one-dimensional approaches in the cases they examined. Both studies originate from technical fields rather than education and therefore do not establish the effectiveness of these methods in vocational assessment. Nevertheless, they reinforce an important methodological principle: function designs need to be tested rather than simply assumed to be appropriate.

The main limitation of the corpus does not lie in the availability of different function types, but in the lack of evidence that the selected parameters produce valid and consistent assessments. Nine of the 13 studies received low ratings for boundary calibration, whereas only Doz et al. (2022) used a large sample of 2,279 participants. The MAPE of 1.35% reported by Izzati et al. (2025) and the system functionality testing reported by Pahtoni and Arifin (2023) cannot be treated as evidence of external validity because both primarily demonstrate the internal performance of their respective systems. In this context, the traceability of each assessment indicator is essential. Sadler (2009) emphasized the integrity of grading in terms of the alignment between assigned grades and the quality of student performance. This means that fuzzy outputs should be able to explain which indicators have been achieved and which remain below the required standards, rather than merely generating a single final score.

Based on these findings, the application of fuzzy logic in vocational practical assessment should adopt a *hybrid crisp-fuzzy* architecture rather than replacing all competency decisions with membership degrees. Mandatory and non-compensatory requirements, such as the use of personal protective equipment, occupational safety procedures, and essential Work Performance Criteria (Kriteria Unjuk Kerja, KUK), should be treated as *crisp gates*. Students who fail to meet these requirements should not be classified as competent merely because they obtain high scores on other indicators. Fuzzy logic should instead be applied to graded dimensions such as precision, time efficiency, neatness, and work independence. Its application requires four conditions: calibration based on empirical data or physical benchmarks, a concise model with three to four indicators, indicator-level diagnostic feedback, and controlled testing against an analytical rubric. This framework is the author's conceptual proposal, not an empirical finding from the corpus.

Conclusion

This review demonstrates that *fuzzy logic* can translate assessment scores into graded categories, but its application to vocational practical assessment cannot automatically be considered appropriate. Although triangular and trapezoidal functions were the most frequently reported forms across the 13 studies, most of the evidence came from academic assessment rather than psychomotor skills measurement. This contextual distinction is important because vocational competencies require observable work procedures, physical performance standards, and safety requirements that cannot be compensated for by scores on other indicators. Therefore, *fuzzy logic* is better positioned as a complement to rubrics rather than a replacement for competency standards. Its application needs to be supported by data-driven membership function calibration, transparent rules, and indicator-level diagnostic feedback. The proposed *hybrid crisp-fuzzy* framework retains mandatory occupational health and safety (OHS) requirements and essential Work Performance Criteria (KUK) as strict competency gates, while using fuzzy logic to assess graded dimensions such as precision, time efficiency, neatness, and work independence.

The contribution of this review lies in mapping fuzzy mechanisms, membership function types, and calibration practices, while identifying methodological requirements for their application in vocational education. It also proposes a *hybrid crisp-fuzzy* framework that preserves mandatory competency requirements while accommodating variations in performance quality. However, the findings should be interpreted in light of the scope and methodological limitations of this *Critical Narrative Review*. Although the literature search covered Scopus, Google Scholar, Garuda, and ERIC, the search process included exploratory stages, and complete search records were not available for all 13 primary studies. Consequently, the selection process cannot be considered fully reproducible, and the possibility of selection bias remains. The evidence base is also limited in its representation of vocational practical assessment: only two studies examined vocational education in Indonesia, and only one directly assessed industrial practice in a vocational high school. Furthermore, external validation remains limited, with no reported controlled comparison between fuzzy-based assessment and analytical rubrics in vocational practical settings. Future research should strengthen the transparency and reproducibility of literature searches, expand empirical investigations using workshop practice data, involve industry assessors and Professional Certification Institutions (LSPs), and conduct controlled comparisons with analytical rubrics to evaluate assessment validity, consistency, and diagnostic value.

References

Abiyyu, A. H., & Widayanti, L. (2023). Sistem pendukung keputusan evaluasi hasil belajar siswa di SMK PGRI 3 Sidoarjo menggunakan metode Fuzzy AHP (Analytical Hierarchy Process). *METHOMIKA: Jurnal Manajemen Informatika &*

Komputerisasi Akuntansi, 7(2), 158–174. https://doi.org/10.46880/jmika.Vol7No2.pp158-174

Adánez, J. M., Al-Hadithi, B. M., & Jiménez, A. (2019). Multidimensional membership functions in T–S fuzzy models for modelling and identification of nonlinear multivariable systems using genetic algorithms. *Applied Soft Computing, 75*, 607–615. https://doi.org/10.1016/j.asoc.2018.11.034

Amelia, N., Abdullah, A. G., & Mulyadi, Y. (2019). Meta-analysis of student performance assessment using fuzzy logic. *Indonesian Journal of Science and Technology, 4*(1), 74–88. https://doi.org/10.17509/ijost.v4i1.15804

Aziz, N. S. B., Rozi, F. F. B., Ghaming, S. N. F. B., & Ridzuan Ishak, N. A. B. (2022). Evaluating students assessment group project using fuzzy analytic hierarchy process (FAHP). *Journal of Mathematics and Computing Science, 8*(2), 1–11. https://doi.org/10.24191/jmcs.v8i2.6965

Bhattacharyya, R., & Mukherjee, S. (2020). Fuzzy membership function evaluation by non-linear regression: An algorithmic approach. *Fuzzy Information and Engineering, 12*(4), 412–434. https://doi.org/10.1080/16168658.2021.1911567

Cui, Y., Guo, Q., Leighton, J. P., & Chu, M.-W. (2020). Log data analysis with ANFIS: A fuzzy neural network approach. *International Journal of Testing, 20*(1), 78–96. https://doi.org/10.1080/15305058.2018.1551225

Doz, D., Felda, D., & Cotič, M. (2022). Assessing students' mathematical knowledge with fuzzy logic. *Education Sciences, 12*(4), Article 266. https://doi.org/10.3390/educsci12040266

Ghozali, F. A., & Haryanto. (2016). Penggunaan logika fuzzy untuk pengembangan instrumen penilaian kinerja praktik sistem kontrol terprogram siswa sekolah menengah kejuruan. *E-Journal Universitas Negeri Yogyakarta, 6*(4), 293–302. https://journal.student.uny.ac.id/index.php/mekatronika/article/download/4233/pdf

Glushkova, T., Ivanova, V., & Zlatanov, B. (2024). Beyond traditional assessment: A fuzzy logic-infused hybrid approach to equitable proficiency evaluation via online practice tests. *Mathematics, 12*(3), Article 371. https://doi.org/10.3390/math12030371

Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal, 26*(2), 91–108. https://doi.org/10.1111/j.1471-1842.2009.00848.x

Hegazi, M. O., Almaslukh, B., & Siddig, K. (2023). A fuzzy model for reasoning and predicting student's academic performance. *Applied Sciences, 13*(8), Article 5140. https://doi.org/10.3390/app13085140

Izzati, A. N., Latif, U. K., & Sintiya, E. S. (2025). Prediction of higher education student academic achievement using Mamdani fuzzy logic. *Journal of Advances in Information and Industrial Technology, 7*(1), 71–82. https://doi.org/10.52435/jaiit.v7i1.682

Ivanova, V., & Zlatanov, B. (2019). Implementation of fuzzy functions aimed at fairer grading of students' tests. *Education Sciences, 9*(3), Article 214. https://doi.org/10.3390/educsci9030214

Mamdani, E. H., & Assilian, S. (1975). An experiment in linguistic synthesis with a fuzzy logic controller. *International Journal of Man-Machine Studies, 7*(1), 1–13. https://doi.org/10.1016/S0020-7373(75)80002-2

Maraj, E., Peposhi, A., & Bendo, A. (2026). Application of fuzzy logic for the evaluation of student academic performance in biomedical subjects. *International Journal of Advances in Applied Sciences, 15*(1), 235–243. https://doi.org/10.11591/ijaas.v15.i1.pp236-244

Namlı, N. A., & Şenkal, O. (2018). Using the fuzzy logic in assessing the programming performance of students. *International Journal of Assessment Tools in Education, 5*(4), 701–712. https://doi.org/10.21449/ijate.429123

Nirawana, I. W. S., Akhyar, R. M., & Sampetoding, E. A. M. (2026). A fuzzy logic approach for student performance and exam integrity assessment in a web-based smart quiz system in higher education. *Information Technology Education Journal, 5*(3), 1–19. https://doi.org/10.59562/intec.v5i3.12904

Özseven, B. E., & Çağman, N. (2022). A novel student performance evaluation model based on fuzzy logic for distance learning. *International Journal of Multidisciplinary Studies and Innovative Technologies, 6*(1), 29–37. https://doi.org/10.36287/ijmsit.6.1.29

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., . . . Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, Article n71. https://doi.org/10.1136/bmj.n71

Pahtoni, T. Y., & Arifin, F. (2023). Design and development of industrial practice monitoring and assessment systems using Tsukamoto fuzzy logic. *Elinvo (Electronics, Informatics, and Vocational Education), 8*(2), 182–192. https://doi.org/10.21831/elinvo.v8i2.57669

Sadler, D. R. (2009). Grade integrity and the representation of academic achievement. *Studies in Higher Education, 34*(7), 807–826. https://doi.org/10.1080/03075070802706553

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research, 104*, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

Suswanto, H., Aripriharta, Handayani, A. N., Mufti, N., Musta'Amal, A. H. B., & Prasetyo, M. W. (2025). A fuzzy-based longitudinal model for evaluating practical learning effectiveness in engineering education. *Cogent Education, 12*(1), Article 2598910. https://doi.org/10.1080/2331186X.2025.2598910

Takagi, T., & Sugeno, M. (1985). Fuzzy identification of systems and its applications to modeling and control. *IEEE Transactions on Systems, Man, and Cybernetics, 15*(1), 116–132. https://doi.org/10.1109/TSMC.1985.6313399

Tang, W.-L., Tsai, J.-T., & Chen, Y.-M. (2021). Fuzzy logic and Gagné learning hierarchy for assessing mathematics skills. *Science Progress, 104*(2). https://doi.org/10.1177/00368504211014346

Thuhra, R., & Kaur, S. (2026). Fuzzy logic for assessing student learning behaviour and academic outcomes: A systematic literature review. *ShodhShreejan: Journal of Creative Research Insights, 3*(1), 102–108. https://doi.org/10.29121/ShodhShreejan.v3.i1.2026.75

Tsai, H.-C., Lee, A.-S., Lee, H.-N., Chen, C.-N., & Liu, Y.-C. (2020). An application of the fuzzy Delphi method and fuzzy AHP on the discussion of training indicators for the regional competition, Taiwan National Skills Competition, in the trade of joinery. *Sustainability, 12*(10), Article 4290. https://doi.org/10.3390/su12104290

Tuan, T. A., Tho, N. D., Nghia, N. H., & Loan, D. T. T. (2025). A systematic review and meta-analysis of fuzzy logic for students' performance assessment. *International Journal of Engineering Trends and Technology, 73*(9), 155–174. https://doi.org/10.14445/22315381/IJETT-V73I9P114

Wei, X. (2024). Performance assessment of college students higher vocational mathematics education using fuzzy evaluation model. *Journal of Intelligent & Fuzzy Systems, 46*(3), 7155–7171. https://doi.org/10.3233/JIFS-235564

Wu, C., Peng, F., Li, C., Jiao, P., & Li, R. (2024). Research on the evaluation of learning quality of higher vocational students based on fuzzy comprehensive evaluation method. In Y. Feng et al. (Eds.), *Proceedings of the 4th International Conference on Internet, Education and Information Technology (IEIT 2024)* (pp. 29–35). Atlantis Press. https://doi.org/10.2991/978-94-6463-574-4_5

Wu, H., Sun, X., Li, D., & Zhao, Y. (2022). Research on learning evaluation of college students based on AHP and fuzzy comprehensive evaluation. *Computational Intelligence and Neuroscience, 2022*, Article 9160695. https://doi.org/10.1155/2022/9160695

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control, 8*(3), 338–353. https://doi.org/10.1016/S0019-9958(65)90241-X

Zangeneh, M., Aghajari, E., & Forouzanfar, M. (2020). A survey: Fuzzify parameters and membership function in electrical applications. *International Journal of Dynamics and Control, 8*(3), 1040–1051. https://doi.org/10.1007/s40435-020-00622-1

Zhang, X., & Wang, A. (2024). Enhancing the performance of vocational education in the digital economy with the application of fuzzy logic algorithm. *International Journal of Computational Intelligence Systems, 17*, Article 185. https://doi.org/10.1007/s44196-024-00591-9