ISSN: 2181-3337 | SCIENTISTS.UZ VOLUME 5 ISSUE 7 JULY 2026

THE MFMR-S QUESTIONNAIRE FOR EVALUATING INDEPENDENT LEARNING ASSIGNMENTS IN A CREDIT-MODULAR SYSTEM: A METHODOLOGICAL EXAMPLE BASED ON SIMULATED DATA

Xolboyev Umrbek
Lecturer, Department of Social Sciences, Gulistan State University
E-mail: xolboyevumr@gmail.com

DOI: 10.67620/sai.57b0019

ABSTRACT
The quality of independent-learning assignments in a credit-modular system cannot be judged from the final product alone; it also depends on the alignment of goal clarity, learning activity, process monitoring, and reflection. This methodological paper presents a preliminary 25-item MFMR-S questionnaire comprising four process dimensions and a separate perceived-effectiveness criterion. The items were derived from constructive alignment, self-regulated learning, formative assessment, and authentic-task literature. To verify scoring and reporting procedures, a fully reproducible synthetic dataset of 50 cases was generated with the fixed random seed 20260805. In the worked example, subscale Cronbach's alpha ranged from 0.771 to 0.861 and one-factor omega from 0.778 to 0.862. Spearman correlations between the process scales and perceived effectiveness ranged from 0.441 to 0.540; the multiple regression model yielded R^2=0.503. These values are not evidence about any student population. They demonstrate coding, diagnostics, and reporting only. MFMR-S should therefore be regarded as a draft instrument rather than a validated scale. Empirical use requires content review, cognitive interviewing, exploratory and confirmatory factor analyses, test–retest assessment, and criterion validation against external learning outcomes.

ACCEPTED

07.08.26

Keywords: *credit-modular system, independent learning, MFMR-S, questionnaire development, simulation, reliability, validation, feedback.*

CC BY

1. Introduction

A credit represents the total amount of work expected of a student to achieve specified outcomes, rather than the number of classroom contact hours in the educational process. The ECTS Users' Guide states that credits should be planned in conjunction with learning outcomes and student workload [2]. The Concept for the Development of Higher Education in Uzbekistan until 2030 likewise identifies the preparation of independently minded specialists with contemporary competencies as one of the system's priorities [1]. An independent assignment should therefore not be treated as a discrete task appended to a course syllabus; it should integrate the intended outcome, learning activity, assessment evidence, and feedback into a single didactic unit.

Constructive alignment requires a coherent relationship among the intended goal, the work performed by the student, and the assessment criteria [3]. The revised version of Bloom's taxonomy helps

move tasks beyond remembering and understanding toward analysis, evaluation, and creation [4]. Models of self-regulated learning conceptualize planning, process monitoring, strategy adjustment, and final self-evaluation as interconnected stages [5, 6, 8]. Formative feedback serves the learning process only when it indicates the learner's next course of action [7, 13, 15].

Short questionnaires designed to evaluate independent learning frequently combine assignment design, student self-regulation, and achieved outcomes within a single scale. Such a mixture may produce high internal consistency, yet it leaves unclear which pedagogical mechanism is being measured. Moreover, predictor and outcome measures obtained from the same respondent, through the same response format, and at the same time point may amplify common method variance [20].

The MFMR-S project is intended to separate these two problems. The acronym MFMR is retained from the source-language labels for four process domains: Goal, Activity, Monitoring, and Reflection. The letter "S" is likewise retained from the source-language label for perceived Effectiveness, which is measured as a separate criterion rather than as an additional process component. The purpose of this article is not to announce a finished, validated scale. Rather, it is to translate theoretical constructs into items, specify reverse-coding and scoring procedures precisely, demonstrate a reproducible analytical sequence using synthetic data, and define a validation program for an empirical study.

2. THEORETICAL BOUNDARIES OF THE MFMR-S MODEL

2.1. Alignment between the Goal and the Learning Outcome

The Goal scale addresses why an assignment is given, which knowledge or skill it is intended to develop, what product is expected, and the criteria by which the work will be assessed. This domain is the practical expression of constructive alignment: the activity performed by the student should provide evidence of the intended learning outcome [3].

2.2. Activity and Assignment Design

The Activity scale was designed to assess whether the task is linked to practical situations, organized in stages, requires higher-order thinking, avoids reliance on copying, and is commensurate with the credit workload. Research on authentic assessment indicates that tasks resembling real professional situations strengthen opportunities to apply and transfer knowledge [14]. In experiential learning, the activity itself, the observations derived from it, and subsequent experimentation are likewise inseparable [16].

2.3. Monitoring and Feedback

Monitoring is not limited to checking the final product. Interim checkpoints, a comprehensible rubric, comments directed toward error correction, and early support for students who are falling behind make the process manageable. The effect of feedback depends more on its specificity, timing, and connection to the next step than on its volume [7, 13, 15].

2.4. Reflection and the Self-Regulation Cycle

The Reflection domain treats planning, self-checking during performance, managing procrastination, adjusting strategy in response to feedback, and drawing conclusions at the end of the work as a single provisional cycle. In the models of Zimmerman and Pintrich, these actions correspond to the phases of forethought, performance control, and final self-evaluation [5, 6]. The models reviewed by Panadero likewise interpret reflection as a mechanism that modifies subsequent activity [8]. The assumption that these five items form a single factor is theoretical and must be tested through factor analysis using empirical data.

2.5. Effectiveness as a Separate Criterion

The Effectiveness scale measures not the assignment process but the student's perception of the learning outcome: deeper understanding of the content, analytical thinking, preparation for professional tasks, perceived influence on learning, and application of knowledge in a subsequent situation. An earlier item concerning self-regulation was removed from the outcome scale because it overlapped conceptually

ISSN: 2181-3337 | SCIENTISTS.UZ

with the Reflection domain. Nevertheless, perceived effectiveness cannot replace an external criterion such as an actual grade, a rubric score, or on-time completion of the assignment.

Table 1.

Operational Map of the MFMR-S Constructs

Construct	Operational content	Items	Theoretical basis and source
Goal	Clarity of the assignment purpose, expected evidence, and assessment criteria	1–5	Constructive alignment [3]
Activity	Practical relevance, sequencing, cognitive demand, and workload alignment	6–10	Bloom's taxonomy; authentic and experiential learning [4, 14, 16]
Monitoring	Interim checks, process monitoring, feedback, and rubric	11–15	Formative assessment and feedback [7, 13, 15]
Reflection	Planning, self-monitoring, revision, and final review	16–20	Self-regulated learning [5, 6, 8]
Effectiveness	Perceived understanding, transfer, and professional preparedness	21–25	Separate outcome criterion [9–12]

3. MATERIALS AND METHODS

3.1. Type of Study and Unit of Analysis

This study represents the preliminary stage of methodological instrument development. No human participants were involved. The unit of analysis was a set of 50 computer-generated synthetic response records. Terms such as "respondent," "student opinion," or "sample result" are not used to describe the synthetic calculations. The numerical outputs demonstrate only how scoring, reliability coefficients, correlations, regression, and diagnostic tables can be produced.

3.2. Item Development and Instrument Status

The initial items were developed deductively from theoretical sources. Five statements were selected for each construct. Responses are recorded on a Likert scale ranging from 1 ("strongly disagree") to 5 ("strongly agree"). To assess automatic acquiescence tendencies, Items 4, 9, 14, 18, and 24 were worded in the reverse direction. Reverse-worded statements may, however, generate a separate method factor; their comprehensibility must therefore be examined in cognitive interviews. The present version is an author-developed draft that has not undergone expert content validation and does not have the status of an applied diagnostic instrument. Scale development should evaluate construct definition, expert appraisal, cognitive testing with the target group, item reduction, factor structure, reliability, and validity as distinct stages [17, 18].

3.3. Generation of Synthetic Data

The reproducible example was generated using the NumPy random-number generator with the seed fixed at 20260805. Four multivariate normally distributed latent variables were first generated for Goal, Activity, Monitoring, and Reflection. Their prespecified correlation matrix is presented in Table 2. The latent Effectiveness variable was constructed as $0.22 \cdot \text{Goal} + 0.38 \cdot \text{Activity} + 0.42 \cdot \text{Monitoring} + 0.30 \cdot \text{Reflection} + 0.62 \cdot \varepsilon$. For each item, a loading between 0.65 and 0.82, item-specific random error, and thresholds of −1.15, −0.35, 0.35, and 1.10 separating five response categories were applied. Reverse-worded statements were transformed in the raw file using $6-x$. The complete code is provided in Appendix 3.

ISSN: 2181-3337 | SCIENTISTS.UZ

Table 2.

Simulation Correlation Matrix for the Four Process Latent Variables

Variable	Goal	Activity	Monitoring	Reflection
Goal	1.00	0.30	0.25	0.25
Activity	0.30	1.00	0.45	0.35
Monitoring	0.25	0.45	1.00	0.40
Reflection	0.25	0.35	0.40	1.00

SHA-256 checksum of the synthetic raw-data CSV file:
fcb71261ceb959fb30ef1afc0bdc52bf2b6d5e1f0c0a25817307e0e9cbd7a22b
This value can be used to verify that re-running the code produces an identical file.

3.4. Coding and Scale Scores

Reverse-worded items were recoded using the formula 6−x. Each subscale score was calculated as the arithmetic mean of its five corresponding items. The synthetic file contains no missing responses. In an empirical study, a prespecified protocol would calculate a subscale mean only when at least four items had been answered; otherwise, the value would be coded as missing. The ranges 1.00–1.80, 1.81–2.60, 2.61–3.40, 3.41–4.20, and 4.21–5.00 are provided solely for descriptive convenience and are not normative cut-offs.

3.5. Statistical Analysis

Descriptive analyses reported the mean, standard deviation, and proportion of responses scored 4 or 5. Internal consistency was expressed using Cronbach's alpha and McDonald's omega derived from a one-factor model. A percentile-bootstrap 95% confidence interval for alpha was calculated from 2,000 resamples. Because alpha depends on restrictive assumptions such as tau-equivalence, it should not be interpreted as the sole source of reliability evidence [19].

Monotonic associations between the process scales and Effectiveness were summarized using Spearman's rho, with 95% bootstrap confidence intervals. In an ordinary least-squares regression in which all four process scores were entered simultaneously as predictors of Effectiveness, the reported statistics were B, standard error, standardized beta, 95% confidence interval, t, p, and VIF. The residual distribution was assessed with the Shapiro–Wilk test; variance stability with the Breusch–Pagan test; linearity with the Ramsey RESET test; autocorrelation with the Durbin–Watson statistic; and influential observations with Cook's distance. P values were retained to display the complete statistical software output; in synthetic data, they do not constitute evidence about a real population.

3.6. Ethical Considerations and Research Integrity

Because no human participants, personal data, or actual academic outcomes were used, participant consent and ethics committee approval were not required for this methodological example. Before empirical testing begins, the relevant ethical approval will be obtained; the voluntary nature of participation will be explained in writing; students will be assured that refusal will not affect their grades; and directly identifying information will be stored separately from scale responses. The author expressly acknowledges the obligation not to present the synthetic results as empirical data.

4. RESULTS OF THE SYNTHETIC WORKED EXAMPLE

The following statistics are products of the data-generation rules described in Section 3.3. They reflect the selected simulation parameters and one random realization of 50 records, rather than the instrument's actual psychometric properties.

ISSN: 2181-3337 | SCIENTISTS.UZ

Table 3.

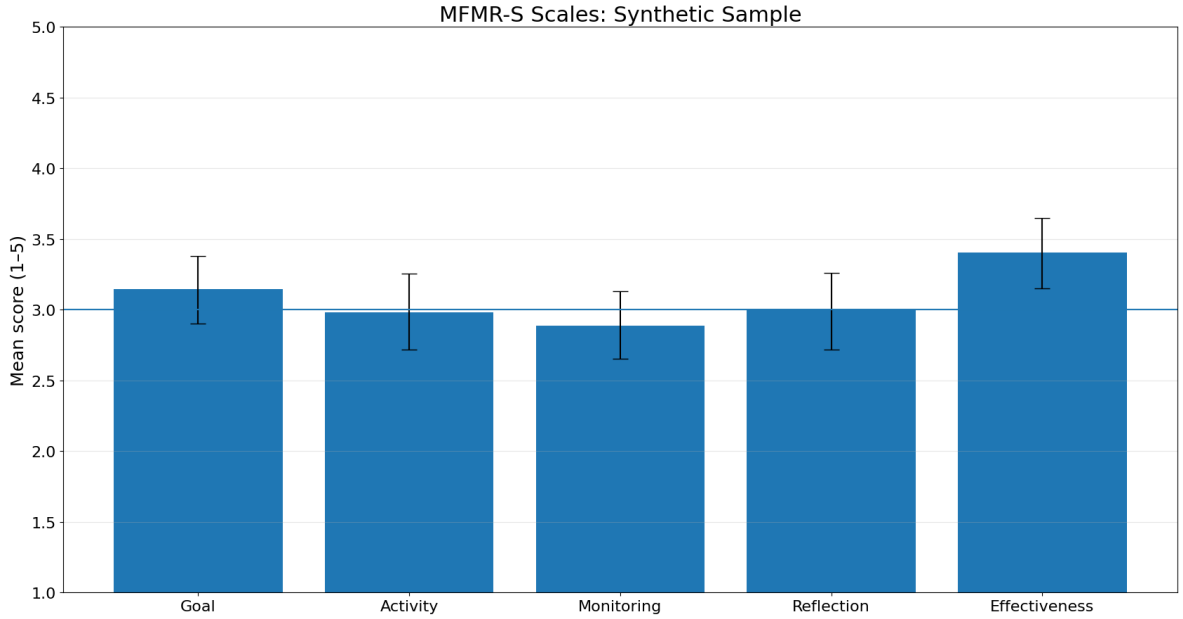
Synthetic Descriptive and Internal-Consistency Statistics for the MFMR-S Subscales (n=50)

Scale	α [95% CI]	ω	M (SD)	4–5, %
Goal	0.771 [0.640; 0.853]	0.778	3.14 (0.85)	39.2
Activity	0.861 [0.774; 0.908]	0.862	2.98 (0.99)	34.0
Monitoring	0.794 [0.651; 0.868]	0.797	2.89 (0.89)	33.6
Reflection	0.851 [0.771; 0.898]	0.854	3.00 (0.99)	38.0
Effectiveness	0.840 [0.746; 0.895]	0.844	3.40 (0.89)	45.2

In the worked example, alpha ranged from 0.771 to 0.861 and omega from 0.778 to 0.862. The proximity of the alpha and omega estimates indicates no marked divergence under the one-factor calculation; however, this pattern does not constitute validity evidence in synthetic data whose factor structure has not been tested. The highest mean was observed for Effectiveness (M=3.40), and the lowest for Monitoring (M=2.89). Because this ordering was shaped in advance through the item thresholds and latent formula, no inference is made about an educational setting.

Figure 1.

Mean MFMR-S Scale Scores and 95% Bootstrap Confidence Intervals



At the item level, the lowest mean was recorded for Q11 (M=2.62), followed by Q10 (M=2.82) and Q12 (M=2.84). The highest values were observed for Q21 (M=3.58), Q23 (M=3.48), and Q25 (M=3.44). Corrected item–total correlations ranged from 0.417 to 0.762. These values do not justify decisions about item retention; with empirical data, polychoric correlations, factor loadings, item information, and content coverage should be considered jointly. The complete item table is provided in Appendix 4.

Table 4.

Spearman Associations between the Process Scales and Effectiveness

Process scale	ρ [bootstrap 95% CI]	p*
Goal	0.540 [0.272; 0.734]	<0.001
Activity	0.441 [0.187; 0.650]	0.001
Monitoring	0.448 [0.169; 0.647]	0.001
Reflection	0.513 [0.257; 0.706]	<0.001

** p values display the output for the synthetic data and are not used for inference to a real population.*

ISSN: 2181-3337 | SCIENTISTS.UZ

All four coefficients were positive and ranged from 0.441 to 0.540. The larger coefficients for Goal and Reflection are specific to this particular simulation realization. Although the latent Effectiveness formula assigned larger coefficients to Monitoring and Activity, the same ordering need not emerge in 50 discretized response records. This discrepancy illustrates why, in simulation studies, a data-generating parameter must not be treated as identical to the coefficient estimated in a sample.

Table 5.

Multiple Regression Predicting Effectiveness: Synthetic Worked Example

Predictor	B (SE)	β	95% CI	t; p*; VIF
Goal	0.210 (0.139)	0.199	[-0.070; 0.490]	t=1.51; p=0.139; VIF=1.57
Activity	0.222 (0.107)	0.246	[0.006; 0.437]	t=2.07; p=0.044; VIF=1.28
Monitoring	0.288 (0.113)	0.285	[0.059; 0.516]	t=2.54; p=0.015; VIF=1.15
Reflection	0.268 (0.111)	0.296	[0.044; 0.491]	t=2.42; p=0.020; VIF=1.36

Model output: R^2=0.503, bootstrap 95% CI [0.346; 0.697], adjusted R^2=0.459, F(4,45)=11.41, p<0.001. Among the standardized coefficients, Reflection (β=0.296) and Monitoring (β=0.285) were larger. The regression interval for Goal included zero; in this synthetic example, this finding demonstrates that a bivariate correlation and a coefficient from a multivariable model do not have the same interpretation.

Table 6.

Regression Diagnostics

Diagnostic domain	Statistic	Result	Interpretation for the synthetic sample
Residual normality	Shapiro–Wilk	W=0.981; p=0.581	No clear departure was detected
Variance stability	Breusch–Pagan	LM=2.59; p=0.629	No indication of heteroscedasticity was observed
Linearity	Ramsey RESET	F=0.17; p=0.686	No requirement for a quadratic specification was indicated
Residual autocorrelation	Durbin–Watson	DW=1.95	The value is close to 2
Influential case	Cook's distance	max D=0.152	Less than 1
Multicollinearity	VIF	1.15–1.57	No serious problem was apparent

5. DISCUSSION

The worked example demonstrates the complete MFMR-S data-processing sequence: recoding reverse-worded items, calculating subscale scores, evaluating internal consistency with two coefficients, distinguishing the relationship between process and outcome, and examining regression assumptions. Nevertheless, none of the calculated coefficients confirms that the instrument is reliable or valid among actual students. The high alpha values are a natural consequence of having specified the simulated item loadings around a single factor in advance. Alpha alone guarantees neither unidimensionality nor adequate content coverage [19].

Separating MFMR from Effectiveness improves construct clarity. The process scales represent instructor-managed assignment characteristics and students' process strategies, whereas Effectiveness represents a perceived outcome. Nevertheless, common method variance remains when all domains are measured in the same questionnaire, at the same time, and with the same response scale [20]. In an empirical study, assignment design could be assessed using an instructor or independent-expert rubric,

ISSN: 2181-3337 | SCIENTISTS.UZ

monitoring could be derived from learning management system logs, reflection could be obtained from student responses, and the outcome could be represented by a grade or a standardized assignment indicator.

The composition of the Reflection scale requires separate examination. Although planning, procrastination, self-monitoring, use of feedback, and final review theoretically belong to one cycle, empirical data may separate them into two or three factors, such as forethought, performance control, and reflection. Factor analysis should therefore be planned not merely to "confirm a five-factor model," but to compare the fit of competing models.

Furthermore, the 50-record sample was not designed to evaluate factor structure. The sample size required for structural models varies substantially according to the number of indicators, factor loadings, number of factors, missing data, and model complexity; a fixed rule specifying a certain number of respondents per item is insufficient [21]. Monte Carlo planning should be used for empirical validation. As a practical initial benchmark, an EFA sample of approximately 200–300 cases may be considered for 25 items [17], but the final sample size should be justified through simulation-based analyses of power and precision.

6. VALIDATION PROGRAM USING EMPIRICAL DATA

Transforming MFMR-S into a publishable measurement instrument requires a sequential, preregistered protocol. A favorable result at one stage does not replace the subsequent stage.

Table 7.

Recommended Validation Stages for MFMR-S

Stage	Participant/source	Primary assessment	Decision criterion
1. Content review	5–7 experts in pedagogy, psychometrics, and credit-modular education	Item–construct alignment, comprehensibility, redundant and missing content	CVI/CVR; revision based on expert comments
2. Cognitive interviews	10–15 students from the target group, in 2–3 rounds	Interpretation of the question, response-selection process, and reverse-worded items	Rewrite or remove ambiguous items
3. Pilot administration	A small group drawn from different years and fields of study	Completion time, missing responses, and floor/ceiling effects	Correct administrative problems
4. EFA and item reduction	Sample size justified by Monte Carlo analysis; practical initial range of 200–300	Polychoric correlations, parallel analysis, and oblique rotation	Evaluate loadings, cross-loadings, content, and reliability jointly
5. Independent CFA	A sample separate from that used for EFA	Fit of five-factor, four-factor, and competing models	CFI/TLI, RMSEA, SRMR, residuals, and theoretical interpretation
6. Reliability	CFA samples and a retest subgroup	Omega, alpha, test–retest reliability, and measurement error	95% intervals and item diagnostics
7. Validity	External grade, rubric, submission status, or standardized test	Convergent, discriminant, and criterion-related validity	Prespecified hypotheses
8. Invariance	Adequately sized groups by sex, year of study, field, or mode of instruction	Configural, metric, and scalar invariance	Test before conducting between-group comparisons

The empirical protocol should be openly preregistered, with exclusion criteria, procedures for handling missing data, and primary and supplementary analyses specified in advance. When scale scores are compared across courses or semesters, mean comparisons should not be made before measurement invariance has been established. Rather than measuring the outcome solely through student self-report, at least one external indicator of learning should be included.

ISSN: 2181-3337 | SCIENTISTS.UZ

7. THE MFMR MAP FOR ASSIGNMENT DESIGN

Until the questionnaire has been validated, the MFMR logic may still be used as a quality-control checklist. In this use, no score is assigned and no diagnosis is made about a student. For each independent assignment, the instructor prepares written answers to five questions: What learning outcome is intended? Through which product will the student demonstrate that outcome? In what stages will the work be completed? When and how will feedback be provided? How will the student analyze the process? This map can be used to check the scope of the assignment against the credit workload, identify interim support points before final grading, and redesign tasks that are susceptible to copying.

Table 8.

Concise MFMR Map for an Independent Assignment

Domain	Quality-control questions
Goal	Is the learning outcome expressed with a precise verb? Are the assessment criteria aligned with the outcome?
Activity	Does the task require student analysis or decision-making? Is its scope appropriate for the time allocated?
Monitoring	Are interim submission points and feedback deadlines specified? Is the rubric provided in advance?
Reflection	Does the student evaluate the plan, errors, strategy, and next steps in writing?
Evidence of effectiveness	What actual product or outcome, beyond perceived benefit, will be obtained?

8. LIMITATIONS

First, all numerical results are synthetic. They do not describe the circumstances of any university, course, subject, or group of students.

Second, the items have not yet undergone expert review or cognitive interviewing with the target group. No conclusion can be drawn regarding content or face validity.

Third, omega was obtained from a one-factor calculation that approximated the data as continuous. For empirical Likert data, polychoric correlations and an ordinal model may be preferable.

Fourth, OLS regression was applied to scale means for 50 synthetic cases. This step demonstrates software output and does not validate the model.

Fifth, Effectiveness is also a self-report measure. Unless it is linked to independent assessment evidence in subsequent research, common method bias will remain.

9. CONCLUSION

MFMR-S is a preliminary questionnaire designed to characterize an independent learning assignment through four process domains—Goal, Activity, Monitoring, and Reflection—and a separate Effectiveness criterion. The reproducible 50-record synthetic example demonstrated the technical format for coding, subscale calculation, internal consistency, correlation, regression, and diagnostic reporting. The numerical results are not empirical evidence and must not be used to make judgments about educational practice. The instrument's scientific status can be established only after content review, cognitive testing, factor validation, reliability assessment, measurement-invariance testing, and examination of associations with external learning outcomes using real participants.

REFERENCES

1. Decree No. PF-5847 of the President of the Republic of Uzbekistan, dated 8 October 2019. "On Approval of the Concept for the Development of the Higher Education System of the Republic of Uzbekistan until 2030."

ISSN: 2181-3337 | SCIENTISTS.UZ

2. European Commission. ECTS Users' Guide 2015. Luxembourg: Publications Office of the European Union; 2015. doi:10.2766/87192.
3. Biggs J, Tang C. Teaching for Quality Learning at University. 4th ed. Maidenhead: Open University Press; 2011.
4. Anderson LW, Krathwohl DR, editors. A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives. New York: Longman; 2001.
5. Zimmerman BJ. Becoming a self-regulated learner: An overview. Theory Into Practice. 2002;41(2):64–70. doi:10.1207/S15430421TIP4102_2.
6. Pintrich PR. A conceptual framework for assessing motivation and self-regulated learning in college students. Educational Psychology Review. 2004;16:385–407. doi:10.1007/s10648-004-0006-x.
7. Nicol DJ, Macfarlane-Dick D. Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. Studies in Higher Education. 2006;31(2):199–218. doi:10.1080/03075070600572090.
8. Panadero E. A review of self-regulated learning: Six models and four directions for research. Frontiers in Psychology. 2017;8:422. doi:10.3389/fpsyg.2017.00422.
9. Broadbent J, Poon WL. Self-regulated learning strategies and academic achievement in online higher education learning environments: A systematic review. The Internet and Higher Education. 2015;27:1–13. doi:10.1016/j.iheduc.2015.04.007.
10. Theobald M. Self-regulated learning training programs enhance university students' academic performance, self-regulated learning strategies, and motivation: A meta-analysis. Contemporary Educational Psychology. 2021;66:101976. doi:10.1016/j.cedpsych.2021.101976.
11. Edisherashvili N, Saks K, Pedaste M, Leijen Ä. Supporting self-regulated learning in distance learning contexts at higher education level: Systematic literature review. Frontiers in Psychology. 2022;12:792422. doi:10.3389/fpsyg.2021.792422.
12. Xu Z, Zhao Y, Zhang B, Liew J, Kogut A. A meta-analysis of the efficacy of self-regulated learning interventions on academic achievement in online and blended environments in K–12 and higher education. Behaviour & Information Technology. 2023;42(16):2911–2931. doi:10.1080/0144929X.2022.2151935.
13. Morris R, Perry T, Wardle L. Formative assessment and feedback for learning in higher education: A systematic review. Review of Education. 2021;9(3):e3292. doi:10.1002/rev3.3292.
14. Vlachopoulos D, Makri A. A systematic literature review on authentic assessment in higher education: Best practices for the development of 21st century skills, and policy considerations. Studies in Educational Evaluation. 2024;83:101425. doi:10.1016/j.stueduc.2024.101425.
15. Wisniewski B, Zierer K, Hattie J. The power of feedback revisited: A meta-analysis of educational feedback research. Frontiers in Psychology. 2020;10:3087. doi:10.3389/fpsyg.2019.03087.
16. Kolb DA. Experiential Learning: Experience as the Source of Learning and Development. Englewood Cliffs, NJ: Prentice Hall; 1984.
17. Boateng GO, Neilands TB, Frongillo EA, Melgar-Quiñonez HR, Young SL. Best practices for developing and validating scales for health, social, and behavioral research: A primer. Frontiers in Public Health. 2018;6:149. doi:10.3389/fpubh.2018.00149.
18. DeVellis RF, Thorpe CT. Scale Development: Theory and Applications. 5th ed. Thousand Oaks, CA: Sage; 2021.
19. McNeish D. Thanks coefficient alpha, we'll take it from here. Psychological Methods. 2018;23(3):412–433. doi:10.1037/met0000144.
20. Podsakoff PM, MacKenzie SB, Lee JY, Podsakoff NP. Common method biases in behavioral research: A critical review of the literature and recommended remedies. Journal of Applied Psychology. 2003;88(5):879–903. doi:10.1037/0021-9010.88.5.879.

ISSN: 2181-3337 | SCIENTISTS.UZ

21. Wolf EJ, Harrington KM, Clark SL, Miller MW. Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. Educational and Psychological Measurement. 2013;73(6):913–934. doi:10.1177/0013164413495237.

APPENDIX 1. PRELIMINARY MFMR-S QUESTIONNAIRE

Instructions. The following statements are preliminary items designed to evaluate how independent learning assignments are organized. The instrument has not been validated; ethical approval and the respondent's informed consent are required before it is used for research purposes. Response options: 1 - strongly disagree; 2 - disagree; 3 - neither agree nor disagree; 4 - agree; 5 - strongly agree. Respondents are not asked to provide their full name, telephone number, or any other directly identifying information.

I. Clarity of the Goal

#	Statement	1	2	3	4	5
1	The purpose of each independent learning assignment is clearly explained to me.	1	2	3	4	5
2	The knowledge, skill, or competency that the assignment is intended to develop is specified.	1	2	3	4	5
3	The performance requirements and expected outcome are communicated before the assignment is given.	1	2	3	4	5
4	What is expected of me in independent assignments is often unclear.	1	2	3	4	5
5	The assignment and its assessment criteria are aligned with the course learning outcomes.	1	2	3	4	5

II. Activity and Assignment Design

#	Statement	1	2	3	4	5
6	The assignments encourage me to apply theoretical knowledge in practical situations.	1	2	3	4	5
7	Independent assignments are structured progressively, from simple to complex.	1	2	3	4	5
8	The tasks require analysis, the use of evidence, and independent decision-making.	1	2	3	4	5
9	Independent assignments are largely limited to copying ready-made information.	1	2	3	4	5
10	The scope of the assignment is commensurate with the time allocated and the credit workload.	1	2	3	4	5

III. Monitoring and Feedback

#	Statement	1	2	3	4	5
11	For major assignments, interim stages and checkpoint dates are specified.	1	2	3	4	5
12	The instructor monitors not only the final outcome of the assignment but also the process of completing it.	1	2	3	4	5
13	Feedback on completed work is provided in a timely manner.	1	2	3	4	5
14	When I make an error, I do not receive sufficient guidance on how to correct it.	1	2	3	4	5
15	The assessment rubric or criteria are provided in advance and are clear to me.	1	2	3	4	5

IV. Reflection and the Self-Regulation Cycle

#	Statement	1	2	3	4	5
16	Before starting an assignment, I prepare a sequence of steps and a time plan.	1	2	3	4	5

ISSN: 2181-3337 | SCIENTISTS.UZ

#	Statement	1	2	3	4	5
17	During the work, I monitor how closely I am following my plan and check my progress.	1	2	3	4	5
18	I often postpone independent assignments until the deadline.	1	2	3	4	5
19	After receiving feedback from the instructor, I reconsider my working method or solution strategy.	1	2	3	4	5
20	At the end of an assignment, I write down what I learned, where I made mistakes, and what I would change next time.	1	2	3	4	5

V. Perceived Effectiveness - Separate Criterion

#	Statement	1	2	3	4	5
21	Independent assignments help me understand the course content more deeply.	1	2	3	4	5
22	Independent learning develops my analytical and critical thinking skills.	1	2	3	4	5
23	The assignments improve my preparedness for future professional tasks.	1	2	3	4	5
24	Independent assignments do not have a substantial effect on my learning outcomes.	1	2	3	4	5
25	I can apply the knowledge acquired through an independent assignment to a subsequent topic or a new situation.	1	2	3	4	5

APPENDIX 2. CODING AND SCORING KEY

Positively keyed items retain their original values. Reverse-worded Items 4, 9, 14, 18, and 24 are recoded using the formula 6−x: 1→5, 2→4, 3→3, 4→2, and 5→1. The mean of the five items is then calculated for each subscale.

Subscale	Items	Formula
Goal	1–5	(Q1+Q2+Q3+Q4R+Q5)/5
Activity	6–10	(Q6+Q7+Q8+Q9R+Q10)/5
Monitoring	11–15	(Q11+Q12+Q13+Q14R+Q15)/5
Reflection	16–20	(Q16+Q17+Q18R+Q19+Q20)/5
Effectiveness	21–25	(Q21+Q22+Q23+Q24R+Q25)/5

Descriptive ranges: 1.00–1.80 - very low; 1.81–2.60 - low; 2.61–3.40 - moderate; 3.41–4.20 - high; 4.21–5.00 - very high. These ranges are not normative. Once empirical data have been collected, they should be reconsidered in light of the distribution, percentiles, measurement error, and the practical consequences of decisions.

APPENDIX 3. CODE FOR REPRODUCING THE SYNTHETIC DATA

The following Python code reproduces the 50×25 raw Likert-response file exactly. NumPy and pandas are required. The expected SHA-256 value for the resulting CSV file is provided in Section 3.3 of the document.

```
import numpy as np
import pandas as pd

seed = 20260805
rng = np.random.default_rng(seed)
n = 50
corr4 = np.array([
 [1.00, .30, .25, .25],
 [.30, 1.00, .45, .35],
 [.25, .45, 1.00, .40],
 [.25, .35, .40, 1.00]])
Z = rng.normal(size=(n, 4)) @ np.linalg.cholesky(corr4).T
G, A, M, R = Z.T
S = .22*G + .38*A + .42*M + .30*R + .62*rng.normal(size=n)
F = np.column_stack([G, A, M, R, S])
F = (F - F.mean(0)) / F.std(0)
```


ISSN: 2181-3337 | SCIENTISTS.UZ

```
loads = {
 'G':[.78,.76,.72,.68,.74], 'A':[.77,.73,.79,.66,.70],
 'M':[.80,.75,.78,.67,.74], 'R':[.74,.77,.65,.72,.79],
 'S':[.82,.78,.74,.67,.76]}
shift = {'G':.10, 'A':.00, 'M':.05, 'R':.02, 'S':.35}
item_shift = {
 'G':[.05,.00,-.05,-.10,.05], 'A':[.05,-.05,.00,-.15,-.20],
 'M':[-.25,-.05,.05,-.10,.05], 'R':[.00,-.15,-.10,.05,.00],
 'S':[.25,.10,.10,-.05,.05]}
cut = np.array([-1.15, -.35, .35, 1.10])
D, q = {}, 1
for fi, key in enumerate(['G','A','M','R','S']):
    for j, lam in enumerate(loads[key]):
        e = rng.normal(size=n)
        y = lam*F[:,fi] + np.sqrt(1-lam**2)*e + shift[key] + item_shift[key][j]
        D[f'Q{q}'] = np.digitize(y, cut) + 1
        q += 1
D = pd.DataFrame(D)
for q in [4, 9, 14, 18, 24]:
    D[f'Q{q}'] = 6 - D[f'Q{q}']
D.to_csv('mfmr_synthetic_seed_20260805.csv', index=False)
```

APPENDIX 4. ITEM-LEVEL SYNTHETIC DIAGNOSTICS

Item	Scale	M (SD)	4–5, %; corrected item–total r
Q1	Goal	3.36 (1.21)	48.0%; 0.647
Q2	Goal	2.98 (1.17)	30.0%; 0.617
Q3	Goal	3.12 (1.14)	40.0%; 0.503
Q4	Goal	3.14 (1.11)	38.0%; 0.542
Q5	Goal	3.12 (1.24)	40.0%; 0.417
Q6	Activity	3.24 (1.27)	48.0%; 0.746
Q7	Activity	2.90 (1.16)	28.0%; 0.616
Q8	Activity	3.00 (1.31)	32.0%; 0.762
Q9	Activity	2.96 (1.21)	30.0%; 0.657
Q10	Activity	2.82 (1.22)	32.0%; 0.612
Q11	Monitoring	2.62 (1.23)	26.0%; 0.618
Q12	Monitoring	2.84 (1.17)	28.0%; 0.660
Q13	Monitoring	3.06 (1.24)	42.0%; 0.594
Q14	Monitoring	3.00 (1.21)	36.0%; 0.454
Q15	Monitoring	2.92 (1.14)	36.0%; 0.552
Q16	Reflection	3.00 (1.29)	38.0%; 0.717
Q17	Reflection	3.00 (1.25)	38.0%; 0.721
Q18	Reflection	2.88 (1.26)	36.0%; 0.510
Q19	Reflection	3.12 (1.19)	40.0%; 0.710
Q20	Reflection	2.98 (1.27)	38.0%; 0.660
Q21	Effectiveness	3.58 (1.20)	48.0%; 0.650
Q22	Effectiveness	3.36 (1.24)	46.0%; 0.558
Q23	Effectiveness	3.48 (1.05)	46.0%; 0.683
Q24	Effectiveness	3.16 (1.20)	42.0%; 0.694
Q25	Effectiveness	3.44 (1.01)	44.0%; 0.652

Note: Reverse-worded items were recoded before the item diagnostics were calculated. The table pertains to the synthetic example and must not be used to make empirical decisions about retaining or removing an item.

ISSN: 2181-3337 | SCIENTISTS.UZ