

BENCHMARKING SATISFACTION WITH A CHEMISTRY MOOC AND BETWEEN-GROUP ANALYSIS AMONG PRESERVICE CHEMISTRY TEACHERS: USING FOCUSED-ISSMRuinan Guo¹*Mohd Shahril Nizam Shaharom¹Kah Heng Chua²

[1] Department of Curriculum and Instructional Technology, Faculty of Education, Universiti Malaya

[2] Department of Mathematics and Science Education, Faculty of Education, Universiti Malaya

*nizamsaril@um.edu.my

ABSTRACT

Learners' satisfaction with online courses is a crucial indicator for evaluating the effectiveness of online instruction. Grounded in DeLone and McLean's (2003) updated Information Systems Success Model (ISSM), this study proposed a satisfaction-focused ISSM by adopting a satisfaction-centric construct reframing and adapting these constructs through operationalizing indicators as satisfaction facets at the measurement level. It further advocated a combined analytical approach that employed the one-sample t test for benchmarking and 2-way ANOVA for examining between-group differences by background factors, thereby enabling multi-stage and longitudinal tracking to evaluate course quality and effectiveness from multiple dimensions of learners' satisfaction and to assess whether an online course is suitable for broad adoption or better targeted to specific background groups. Accordingly, using the Instructional Design for Secondary School Chemistry MOOC as a quantitative case, this study administered multiple online questionnaires to 140 junior preservice chemistry teachers, who were stratified by two crossed background factors at a representative normal university in Henan Province, China, and supplemented these data with backend system logs on their learning behavior and exam results. The findings indicated consistently favorable satisfaction evaluations at each stage of learning, suggesting value for enhancing chemistry instructional design abilities and promising prospects for broad implementation.

Keywords: *Satisfaction, ISSM, MOOC, analytical methods, preservice chemistry teacher.*

INTRODUCTION

In recent years, online courses have become a major educational pathway alongside face-to-face instruction. With the rapid development of online education, learner satisfaction has gradually become one of the key indicators for evaluating course quality and instructional effectiveness (Yin et al., 2024). Mathew et al. (2024) indicated that course satisfaction is influenced by multiple factors such as instructional quality, course design, and service feedback, while Rahmani et al. (2024) and Dai et al. (2024) found that course satisfaction is associated with learning indicators including completion rates and the intention to continue using the course. Satisfaction research grounded in benchmark-based evaluation can specify levels of perceived satisfaction and guide improvement (Chiu & Chen, 2024). Satisfaction is, in essence, a cognitive and affective response that evolves over time and is subject to long-term influences rather than being static (Anić & Tončić, 2024; Yu, 2022). However, most studies rely on short learning phases or immediate or end-of-unit feedback (Martin & Bolliger, 2022; Yu, 2022; Zeng & Wang, 2021). Measuring only short-term data makes it difficult to capture the dynamic processes of learners' genuine emotions and attitudes (Yu, 2022). Additionally, existing research remains weak in

systematized frameworks dedicated specifically to satisfaction, particularly lacking models or systems that can operate reliably in long-term learning contexts (Martin & Bolliger, 2022; Mustafa & Ali, 2023). Current studies often borrow generic information-systems and technology-acceptance frameworks as macro explanatory scaffolds; however, their applications typically take “overall effectiveness” or “adoption or continuance” as the central axis; satisfaction is thus frequently relegated to a subordinate attribute or mediating position, and the selection of indicators and the organization of structural levels lack a clear operationalization pathway that takes satisfaction as the core construct (e.g., Hamad et al., 2024; Li et al., 2022). From a methodological perspective, empirical investigations of online-course satisfaction are typically conducted through comprehensive post-course questionnaire surveys (Berlin & Weavera, 2022; Lu & Smiles, 2024; Reichgelt & Smith, 2024), sometimes supplemented by participant interviews as qualitative additions (Cheng et al., 2023; Tian & Lu, 2022). Stage-specific fluctuations and reappraisals in learner experience are usually compressed into a single-time-point overall measurement of satisfaction or a one-off cross-sectional comparison, which makes it difficult to capture differences in experience across different stages and internal changes in the structure of satisfaction (Elkins et al., 2021; Schriek et al., 2024).

The Information Systems Success Model (ISSM) was initially proposed by DeLone and McLean (1992), and in 2003 it underwent a major revision to better align with the need to evaluate success in increasingly complex information systems environments (DeLone & McLean, 2003). In its classical applications, the ISSM likewise emphasizes a holistic assessment of the success and value of information systems (Sabeh et al., 2021), with its multidimensional constructs serving as a macro-level reference framework (e.g., Alfaki, 2021; Elmunsyah et al., 2023). This study's innovation lay in a focused application of the updated ISSM, tailoring and specializing the model into a “satisfaction-centered, hierarchically layered” adaptation. By introducing multi-stage satisfaction evaluations, this focused framework can provide dynamic, multidimensional feedback, thus correcting the limitations of one-off cross-sectional aggregation.

One-sample t test is used to compare the sample mean against an explicit benchmark, providing statistical evidence and confidence intervals for whether the target is “met”, thereby facilitating actionable pass or fail determinations and quality monitoring, and it can be paired with effect size to assess the magnitude of improvement and practical significance (Ponce-Renova, 2022). In contrast, 2-way ANOVA can simultaneously identify the main effects and interaction of two background factors, revealing differences among learning groups and which combinations benefit the most; on this basis, one can judge whether to implement stratified dissemination or to optimize the allocation of interventions (Andrade, 2024; Sureiman & Mangera, 2020). In this study, a joint quantitative approach was adopted, combining a one-sample t test, followed by a two-way ANOVA, to evaluate satisfactions. Taken together, with intentions, behaviors and scores as auxiliary support, this approach could assess whether an online course is suitable for widespread adoption or is better oriented toward a specific background group, thereby providing a basis for judging and improving (Dong et al., 2023; Pham et al., 2024; Tipton & Olsen, 2022). While neither method is novel in itself, their combined use forms an integrated, complementary chain of evidence from “whether the benchmark is met” to “for whom satisfaction is higher”, offering greater systematization and explanatory power than only a single test.

This study selected Massive Open Online Course (MOOC) as a representative form of online courses because its large scale, openness, and trackable structure allow complete behavioral logs and stage-based outputs to accumulate on the same platform (Zhu et al., 2022). Specifically, this study adopted the Instructional Design for Secondary School Chemistry MOOC (Liu, 2024) on the Chinese University MOOC platform as a quantitative case. It had a long duration and an integrated task chain, making it suitable for stage-wise tracking of changes, and the participating student cohort was both representative and diverse in background. Its training objectives for preservice chemistry teachers were also highly aligned with the teaching standards issued by the Ministry of Education of the People's Republic of China (2020, 2021, 2022), thus constituting a practice-oriented and transferable real-world empirical case. Accordingly, the research objective of this study was to use the satisfaction-focused ISSM and the joint analytical approach to benchmark preservice chemistry teachers' satisfaction with different aspects of

the Instructional Design for Secondary School Chemistry MOOC, and to further analyze between-group differences across key background factors.

Research Questions

1. What were the levels of satisfaction among preservice chemistry teachers with:
 - a. The system quality
 - b. The information quality
 - c. The service qualityin the experience of the Instructional Design for Secondary School Chemistry MOOC?
2. What were the levels of satisfaction among preservice chemistry teachers with:
 - a. Their experience in the Instructional Design for Secondary School Chemistry MOOC?
 - b. The overall benefits across the entire learning process in the Instructional Design for Secondary School Chemistry MOOC?
 - c. The whole of the Instructional Design for Secondary School Chemistry MOOC?
3. Did the satisfaction levels in RQ1–RQ2 differ by:
 - a. National Teacher Certification Examination status?
 - b. Previous online learning experience?
 - c. Their interaction?

METHODOLOGY

Focused-ISSM for Satisfaction

In DeLone and McLean's ten-year update, it was explicitly stated that the model can be contextually adjusted and flexibly applied to address the measurement challenges of the new e-commerce world (DeLone & McLean, 2003). Building on their updated ISSM, and without altering the boundaries of its original constructs, the interrelationships among constructs, or the integrity of its complete paths, this study adopted a satisfaction-centric construct reframing for the constructs of "system quality", "information quality", "service quality", "user satisfaction", and "net benefits"; at the measurement level, these constructs were adapted by operationalizing the indicators as satisfaction facets. This focused-ISSM is therefore highly suitable for the specialized exploration and evaluation of satisfaction in online courses, particularly long-term, multi-stage online courses. The overall process is explained as follows: after students complete the initial experience phase of the online course, satisfaction with "system quality", "information quality", and "service quality", as well as "user satisfaction with experience" were measured. In addition, it measures students' "intention to continue using" to judge whether the course can proceed smoothly. Thereafter, students continue to complete the online course in full; upon course completion, it measures their "satisfaction with the overall benefits" and "overall satisfaction with the course", and again measures students' "intention to review". Furthermore, throughout the entire learning process, students' "learning behavior"—which corresponds to the original "use" construct—and the "exam results" at each stage—which correspond, in another part, to the original "net benefits" construct—are systematically recorded. Together with the two times measurements of "intention", these serve as auxiliary evidence supporting the "satisfactions" results. The foregoing constitutes the complete procedure.

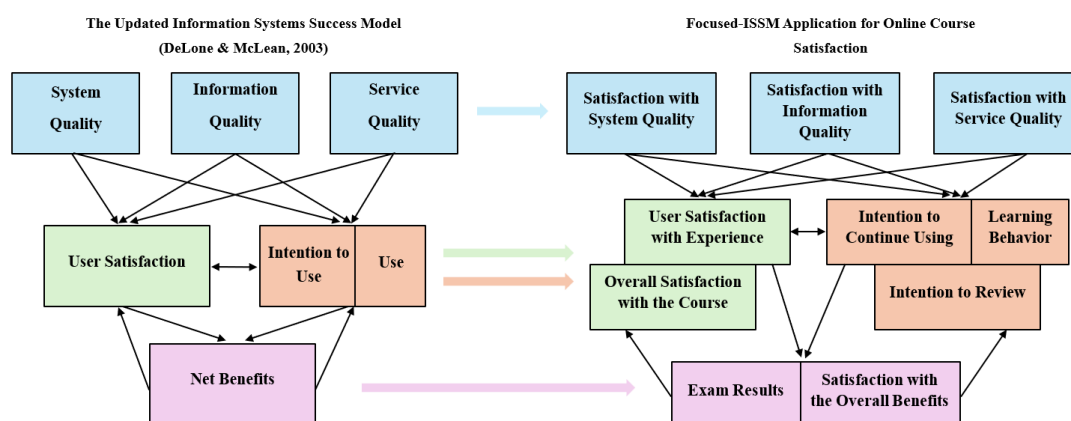


Figure 1. Interpretation for the Satisfaction-Focused Updated ISSM

Figure 1 details how the constructs in the original 2003 D&M updated ISSM shown in the left panel are correspondingly reframed and interpreted at the indicator-level operationalization within the “satisfaction-focused” application framework shown in the right panel.

Description of the Instructional Design for Secondary School Chemistry MOOC

The Instructional Design for Secondary School Chemistry MOOC was designed and instructed by Professor Liu from a normal university in Henan Province, China, and has been officially launched on the Chinese University MOOC platform. This study was based on the data from the 17th session of this MOOC. It lasted 10 weeks, covered 10 major chapters and 52 sub-sections, was divided into three chapters stages, and included three stage tests and a final exam. Detailed contents and arrangements were scheduled and provided in Figure 2.

Chapters Stages	Week	Chapter No.	Courses Content		
Chapters Stage 1	Week 1	Chapter 1- Overview of Chemistry Instructional Design	The Concept and Development of Instructional Design		
			The Levels, Process, and Significance of Instructional Design		
			Theoretical Foundations of Chemistry Instructional Design (I)		
	Week 2	Chapter 2- Background Analysis of Chemistry Instructional Design	Learning Needs Analysis		
			Analysis of Chemistry Instructional Content		
			Analysis of Chemistry Learners		
	Week 3	Chapter 3- Design of Chemistry Instructional Objectives	The Concept and Functions of Instructional Objectives		
			Classification of Instructional Objectives		
			The Construction of a Chemistry Instructional Objectives Framework		
			Design Principles for Chemistry Instructional Objectives		
Chapters Stage 2	Week 4	Chapter 4- Design of Chemistry Instructional Strategies	Formulation of Chemistry Instructional Objectives		
			From Content Standards to Classroom Instructional Objectives		
			The Design Process of Chemistry Instructional Objectives		
			Development of Integrated Instructional Objectives		
			Core Competencies in Chemistry Education		
			Design of Competency-Based Instructional Objectives and Case Studies in Chemistry		
			Chapters Stage 1 Test for Chapters 1 to 3 (In Week 3)		
			Overview of Instructional Strategies		
			Overview of Chemistry Instructional Strategies		
			Instructional Strategies for Teaching Factual Knowledge in Chemistry		
Instructional Strategies for Teaching Theoretical Knowledge in Chemistry					
Instructional Strategies for Teaching Procedural Knowledge in Chemistry					
Instructional Strategies for Affective Learning in Chemistry					
Inquiry-Based Instructional Strategies in Chemistry					
Instructional Strategies for Self-Directed and Collaborative Learning in Chemistry					
Conceptual Change-Oriented Instructional Strategies in Chemistry					
Chapters Stage 3	Week 5	Chapter 5- Design of Chemistry Instructional Scenarios	Instructional Scenarios in Chemistry		
			Principles for Designing Instructional Scenarios in Chemistry		
	Week 6	Chapter 6- Chemistry Experiments and Experimental Design	Methods for Designing Chemistry Instructional Scenarios (I)		
			Methods for Designing Chemistry Instructional Scenarios (II)		
			Characteristics of High-Quality Instructional Scenarios in Chemistry		
			Overview of Chemistry Experiments		
	Week 7	Chapter 7- Design of Chemistry Instructional Assessment	Educational Functions of Chemistry Experiments		
			Chemistry Experiments Design		
	Week 8	Chapter 8- Design of Chemistry Instructional Plan	Control of Experimental Conditions in Chemistry		
			Trends in Chemistry Experiments Design		
Week 9	Chapter 9- Practicing and Reflecting on Chemistry Instructional Design	Instructional Models in Chemistry Experiments			
		Functions, Types, and Principles of Chemistry Instructional Assessment			
Week 10	Chapter 10- Chemistry Lesson Presentation	Design of Chemistry Instructional Assessment			
		Development of Chemistry Example Problems and Practice Exercises			
		Design of Chemistry Test Papers			
		Designing the Process of Chemistry Instruction			
Chapters Stage 2 Test for Chapters 4 to 6 (In Week 6)					
Development of Chemistry Lesson Plans					
Design of Chemistry Student Learning Sheet					
Implementation of Chemistry Instructional Design					
Reflection on Chemistry Instructional Design					
Case Study 1: Implementation of Chemistry Instructional Design					
Case Study 2: Implementation of Chemistry Instructional Design					
The Concept and Purpose of Chemistry Lesson Presentation					
Content of Chemistry Lesson Presentation					
Skills for Effective Chemistry Lesson Presentation					
Case Study 1: Chemistry Lesson Presentation					
Case Study 2: Chemistry Lesson Presentation					
Chapters Stage 3 Test for Chapters 7 to 10 (In Week 10)					
Final Exam (In Week 10)					

Figure 2. Schedule of the Instructional Design for Secondary School Chemistry MOOC

The preservice chemistry teachers needed to follow the weekly study schedule and tasks, completing the self-paced learning of the corresponding chapters’ MOOC’s videos under the “Courseware” module on the platform. They were also required to participate in the discussions and interactions on questions posted by the instructor in the “Discussion Forum” module. As the MOOC progressed through the

designated weeks (chapters stages), they must complete the chapters stage tests in the "Quizzes and Assignments" module. Upon completing the full ten-week duration of the MOOC, they must take the final exam in the "Take an Exam" module before the deadline set by the instructor. Figure 3 below provided a concise overview of the MOOC's interface.

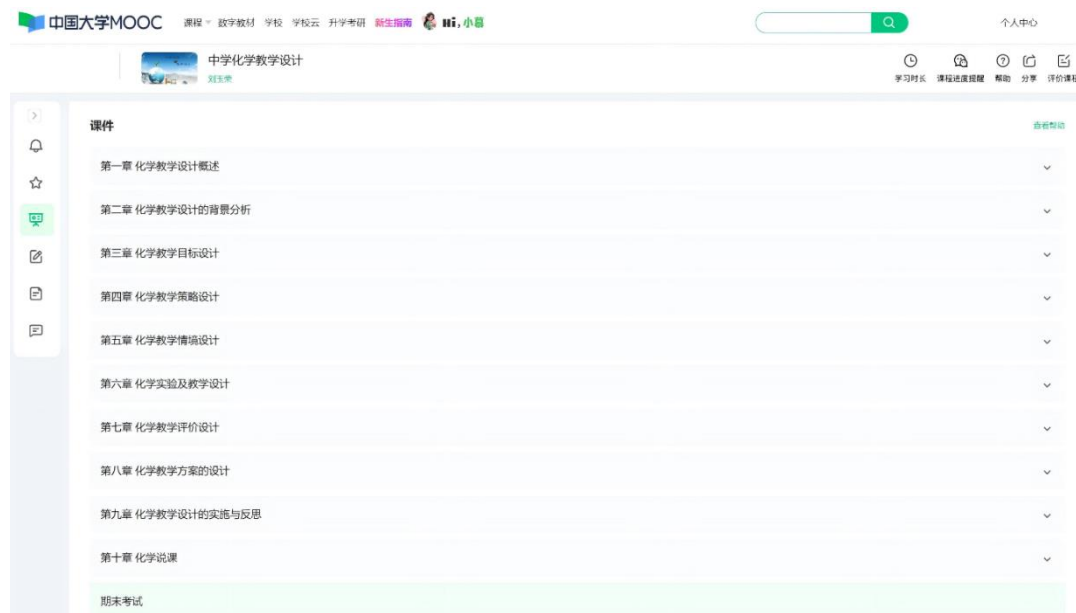


Figure 3. Interface of the Instructional Design for Secondary School Chemistry MOOC

The experience period of preservice chemistry teachers with the MOOC in this study referred to chapters stage 1 (weeks 1, 2 and 3). This was because the chapters stage 1 already includes a comprehensive range of activity types, which was highly representative. The subsequent two chapters stages were structurally consistent with the chapters stage 1, and participants' learning basically followed very similar patterns.

Population and Samples

This study was conducted at the School of Chemistry and Chemical Engineering in a normal university located in Xinxiang, Henan Province, China, where the MOOC instructor was employed and which was considered reasonably representative among normal universities nationwide. The target group was third-year undergraduate students majoring in chemistry education, totaling 216 students, because the MOOC was scheduled in the curriculum for this year level and the cohort also allowed for diversity in participants' learning backgrounds. This study employed the Krejcie and Morgan (1970) formula ($\chi^2 = 3.841$, 95% confidence level; $N = 216$; $P = 0.50$; $d = 0.05$) to determine a required sample size of 140 and used stratified random sampling based on two key grouping variables—whether the preservice chemistry teachers have passed the National Teacher Certification Examination or not and have or no online learning experience. The rationale for selecting these two background factors for the difference analyses was that the former captures variation in preservice teachers' prior knowledge base (Luik & Taimalu, 2021) and teacher professional identity (Lai & Jin, 2021), which may affect their judgments about the usefulness of the course content; the latter may influence learners' expectations regarding the platform's usability and value (Zhang et al., 2023). However, an important issue in the stratification was the overlap between the two backgrounds, which yielded four cross-classified groups (see Figure 4).

	Whether or not have passed the National Teacher Certification Examination	
Online learning experiences	Have passed National Teacher Certification Examination and have online learning experience	Have not passed National Teacher Certification Examination but have online learning experience
	Have passed National Teacher Certification Examination but no online learning experience	Have not passed National Teacher Certification Examination and no online learning experience

Figure 4. Schematic Diagram of the Four Cross-Combined Groups Formed

Therefore, this study further adopted stratified nested random sampling. This ensured adequate representation of all relevant student types. Based on the demographic data from questionnaire, the proportions of each of the four combinations in the population were calculated. However, it was difficult to ensure approximately equal numbers in each combination, which might lead to unequal sample sizes in subsequent analyses. Therefore, the sample size for each combination was adjusted and fixed at 35. In data analyses, if necessary, weights would be applied to correct group-size bias introduced by this adjustment. This would restore the true group proportions, ensuring that the results more accurately reflected the population. If the adjusted sample size for each group closely resembled its distribution in the original population, weighting would not be applied to avoid unnecessary complexity.

Data Collection and Data Analysis

This study collected quantitative data using a series of questionnaires through the web-based survey platform Sojump (WJX), with 5-point Likert-scale items for investigating satisfactions and intentions, and retrieving participants’ learning behaviors and exam results from the backend system logs.

All questionnaires underwent expert review for content validity and were pilot-tested. All Cronbach’s α coefficients ranged from 0.70 to 0.85, indicating good internal consistency. The KMO values exceeded 0.70, and the Bartlett’s Test of Sphericity were significant ($p < 0.001$). EFA supported a unidimensional structure (all loadings > 0.60 ; cumulative variance explained 50–60%), and CFA indices indicated good fit, confirming structural validity. Different questionnaires were administered at different stages. The Personal Information Questionnaire was distributed to the population prior to sampling to collect demographic data, which served as the basis for determining the sample. After the experience period, participants completed the Satisfaction with Qualities in the MOOC Questionnaire, the Satisfaction with Experience Questionnaire, and the Intention to Continue Using Questionnaire online; after completing the entire course content, they further completed the Satisfaction with the Overall Benefits across the Entire Learning Process Questionnaire, the Overall Satisfaction with the Whole of the MOOC Questionnaire, and the Intention to Review Questionnaire.

The learning behavior referred to their frequency of videos learning (at least 45 times in total) and forum interactions (at least 30 times in total) across the MOOC. The exam results referred to the composite “Comprehensive Final Score” (final exam score 50% + three stage-test scores in total 35% + forum interaction score 15%), with performance levels set as passing (≥ 60) and excellent (≥ 85). The criteria and scoring rule were determined by the instructor, and the corresponding records were automatically logged and calculated throughout participants’ learning process.

All satisfaction measures were subjected to inferential statistical analyses using a two-step procedure, with one-sample t tests (two-tailed, $\alpha = 0.05$) first conducted for benchmarking purposes, followed by 2-way ANOVA to examine between-group differences, whereas intentions at two time points, learning behavior, and exam results—serving as auxiliary supporting data—were examined only with one-sample

t tests to determine whether they met the benchmark. In addition to reporting p values, Cohen's d was also reported for the t test, and partial eta squared (η_p^2) was reported for the 2-way ANOVA, to provide effect-size evidence and to strengthen the practical interpretability of the results. According to Goss-Sampson (2022), the one-sample t test assumes approximate normality of the sample distribution, while 2-way ANOVA primarily assumes homogeneity of variances and normality of residuals. However, when sample sizes are large and group sizes are roughly balanced, these tests typically exhibit relative robustness to deviations from normality due to the central limit theorem (Zygmunt, 2023). Therefore, this combined approach is particularly well suited to the analysis of large-enrollment online courses in naturalistic instructional settings.

In educational research, data analysis based on a 5-point Likert scale typically defines different levels according to mean-score ranges (Alkharusi, 2022). Table 1 below demonstrated the interpreted mean-score ranges for satisfaction and intention, which served as interpretive benchmarks for classifying the results.

Table 1. *Interpretive Benchmarks for Satisfaction and Intention: Mean-Score Ranges*

The Mean-Score Range	Satisfaction Level	Intention Level
4.50–5.00	Very High	Extremely Strong
3.50–4.49	High	Strong
2.50–3.49	Moderate	Moderate
1.50–2.49	Low	Weak
1.00–1.49	Very Low	Extremely Weak

Typically, when the level of mean satisfaction score reaches "high" level, it indicates that learners approve of the learning process, suggesting that the learning process has met learners' basic expectations and is capable of providing a relatively good learning experience and outcomes (Zhou, 2024). Hence, the one-sample t tests initially used $\mu_0 = 3.50$ as the benchmark for the "high" level. When the mean difference indicated a substantial margin above 3.50, a supplementary test with $\mu_0 = 4.50$ was performed to evaluate whether the "very high" benchmark was met.

FINDINGS

In all analyses below, to ensure clarity and conciseness, the names of the four cross-combined groups were abbreviated using simple codes. The "National Teacher Certification Examination" was denoted as "N" (sometimes also "NTCE"), and "Online Learning Experience" as "O" (sometimes also "OnlineExp"). "Have passed" or "have" were coded as the character "1", while "have not passed" or "no" as "0". Additionally, all Cohen's d values for one-sample t tests used the sample standard deviation, and all Levene's tests tested the null hypothesis of homogeneity of error variances across groups. The superscript "a" identified the dependent variable for each test, and the superscript "b" indicated the design was "Intercept (omitted from display) + NTCE + OnlineExp + NTCE * OnlineExp". Therefore, footnotes in the following tables were omitted, as all analyses conformed to a consistent pattern.

Demographics

Table 2. Demographics of the Four Cross-Combined Groups

Population = 216, Total Sample Size (n) = 140						
Group	Number of People					Average Age
	In Population	In Sample (Calculated)	In Sample (Adjusted)	Male	Female	
N1O1	57	37	35	17	18	20.8
N0O1	53	34	35	16	19	21.1
N1O0	55	36	35	18	17	21.5
N0O0	51	33	35	16	19	21.7

According to Table 2, since the adjusted sample size closely matched the sample size calculated based on actual proportions, no further weighting was necessary, and the sample size after adjustment remains valid for maintaining the authenticity of the overall data. Moreover, the participants' gender distribution and average age in different groups demonstrated a high degree of balance, which effectively avoided potential biases due to age differences and gender imbalances.

Normality checks were conducted for each variable's data or residuals. All distributions were approximately normal or exhibited only slight skewness and very few mild outliers. Given the relatively large sample sizes across groups, analytic results' robustness could be assured.

Satisfaction with the System Quality

Table 3. One-Sample T Test and Effect Sizes for Mean Satisfaction Scores with the System Quality

Test Value = 3.50									
Groups	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference		Cohen's d Point Estimate	95% Confidence Interval	
					Lower	Upper		Lower	Upper
N1O1	15.705	34	< .001	.92	.80	1.04	2.655	1.940	3.360
N0O1	14.374	34	< .001	.94	.80	1.07	2.430	1.761	3.088
N1O0	9.884	34	< .001	.66	.52	.79	1.671	1.149	2.181
N0O0	9.192	34	< .001	.69	.54	.85	1.554	1.053	2.043
All	23.104	139	< .001	.80	.73	.87	1.953	1.668	2.234

Table 3 showed that all four groups scored significantly above the benchmark of 3.50 (all $t(34) \geq 9.192$, all $p < 0.001$). Mean differences (MDs) with 95% CIs were largest for the O1 groups (N1O1: 0.92 [0.80, 1.04]; N0O1: 0.94 [0.80, 1.07]) and smaller for the O0 groups (N1O0: 0.66 [0.52, 0.79]; N0O0: 0.69 [0.54, 0.85]). For the overall sample, all participants also exceeded the 3.50 benchmark: $t(139) = 23.104$, $p < 0.001$, and the MD was 0.80 (95% CI [0.73, 0.87]). Furthermore, Cohen's d values with 95% CIs were N1O1 = 2.655 [1.940, 3.360], N0O1 = 2.430 [1.761, 3.088], N1O0 = 1.671 [1.149, 2.181], N0O0 = 1.554 [1.053, 2.043] and overall = 1.953 [1.668, 2.234], all exceeding 0.8 and indicating large effect sizes. Therefore, all groups and the overall sample had significantly reached the "high" satisfaction level.

Table 4. *Levene’s Test and 2-Way ANOVA for Mean Satisfaction Scores with the System Quality of All Participants (n = 140)*

Levene’s Test of Equality of Error Variances ^{a,b}					
	Levene Statistic	df1	df2	Sig.	
Based on Mean	.748	3	136	.525	
Based on trimmed mean	.771	3	136	.512	
2-Way ANOVA					
Descriptive Statistics-Mean	N1	N0	O1	O0	
	4.29	4.31	4.43	4.18	
Source	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	3	.759	4.877	.003	.097
NTCE	1	.022	.141	.708	.001
OnlineExp	1	2.250	14.464	< .001	.096
NTCE * OnlineExp	1	.004	.026	.873	.000

Table 4 firstly reported the results of Levene’s test among all participants. Whether based on the mean ($F = 0.748$, $df_1 = 3$, $df_2 = 136$, $p = 0.525 > 0.05$), or based on the trimmed mean ($F = 0.771$, $df_1 = 3$, $df_2 = 136$, $p = 0.512 > 0.05$), the null hypothesis of equal error variances could not be rejected. Furthermore, the 2-way ANOVA (corrected model) was significant ($df = 3$, $F = 4.877$, $p = 0.003 < 0.05$), and its $\eta_p^2 = 0.097 > 0.06$ indicated a medium effect size approaching the 0.14 of large effect size. Among individual effects, “OnlineExp” has a significant effect ($df = 1$, $F = 14.464$, $p < 0.001$, $\eta_p^2 = 0.096$), indicating that preservice chemistry teachers whether or not have online learning experience had a substantial influence on their satisfaction with the system quality in the experience of this MOOC. Mean satisfaction was higher for O1 than O0 ($4.43 > 4.18$; $\Delta = 0.25$), but both groups were at the “high” level. In contrast, NTCE ($df = 1$, $F = 0.141$, $p = 0.708 > 0.05$, $m_{N1} = 4.29$, $m_{N0} = 4.31$, $\eta_p^2 = 0.001$) and their interaction effect ($df = 1$, $F = 0.026$, $p = 0.873 > 0.05$, $\eta_p^2 \approx 0.000$) did not have the significant impact. Collectively, prior online learning experience was the primary driver of between-group differences in system-quality satisfaction.

Satisfaction with the Information Quality

Table 5. *One-Sample T Test and Effect Sizes for Mean Satisfaction Scores with the Information Quality*

Test Value = 3.50									
Groups	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference		Cohen’s d Point Estimate	95% Confidence Interval	
					Lower	Upper		Lower	Upper
N1O1	14.224	34	< .001	.93	.80	1.06	2.404	1.741	3.058
N0O1	9.434	34	< .001	.71	.55	.86	1.595	1.087	2.092
N1O0	16.241	34	< .001	.94	.82	1.06	2.745	2.011	3.470
N0O0	9.535	34	< .001	.63	.49	.76	1.612	1.101	2.112
All	23.042	139	< .001	.80	.73	.87	1.947	1.664	2.228

Table 5 showed that all four groups scored significantly above the benchmark of 3.50 (all $t(34) \geq 9.434$, all $p < 0.001$), with all MDs (95% CIs) entirely above zero. For the overall sample, all participants yielded a result ($t(139) = 23.042$, $p < 0.001$), and the MD was 0.80 (95% CI [0.73, 0.87]). Moreover, all effect

sizes exceeded 0.8, indicating large effects. Thus, all groups and the overall sample had significantly reached the "high" satisfaction level.

Table 6. *Levene's Test and 2-Way ANOVA for Mean Satisfaction Scores with the Information Quality of All Participants*

Levene's Test of Equality of Error Variances ^{a,b}					
	Levene Statistic	df1	df2	Sig.	
Based on Mean	.636	3	136	.593	
Based on trimmed mean	.668	3	136	.573	
2-Way ANOVA					
Descriptive Statistics-Mean	N1	N0	O1	O0	
	4.44	4.17	4.32	4.29	
Source	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	3	.874	5.678	.001	.111
NTCE	1	2.511	16.310	< .001	.107
OnlineExp	1	.036	.235	.629	.002
NTCE * OnlineExp	1	.075	.490	.485	.004

Table 6 first revealed the null hypothesis of equal error variances could not be rejected based on Levene's test among all participants (all $p > 0.05$). Further, the 2-way ANOVA (corrected model) was significant ($p = 0.001 < 0.05$), and its $\eta_p^2 = 0.111 > 0.06$ indicated a medium effect size approaching the 0.14 of large effect size. "NTCE" has a significant effect ($p < 0.001$, $\eta_p^2 = 0.107$) suggesting that preservice chemistry teachers whether or not have passed the NTCE had a substantial influence on their satisfaction with the information quality in the experience of this MOOC. Mean satisfaction was higher for N1 than N0 (4.44 > 4.17; $\Delta = 0.27$), but both groups were at the "high" level. Conversely, online learning experience ($p = 0.629 > 0.05$) and their interaction effect ($p = 0.485 > 0.05$) did not have the significant impact. Altogether, whether or not have passed the NTCE was the primary driver of between-group differences in information-quality satisfaction.

Satisfaction with the Service Quality

Table 7. *One-Sample T Test and Effect Sizes for Mean Satisfaction Scores with the Service Quality*
Test Value = 3.50

Groups	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference		Cohen's d Point Estimate	95% Confidence Interval	
					Lower	Upper		Lower	Upper
					N1O1	16.241		34	< .001
N0O1	13.886	34	< .001	.93	.79	1.06	2.347	1.695	2.989
N1O0	11.950	34	< .001	.76	.63	.89	2.020	1.433	2.596
N0O0	10.030	34	< .001	.68	.54	.82	1.695	1.170	2.211
All	24.994	139	< .001	.83	.76	.89	2.112	1.813	2.409

Table 7 presented that all four groups scored significantly above the benchmark of 3.50 (all $t(34) \geq 10.030$, all $p < 0.001$), with all MDs (95% CIs) entirely above zero. For the overall sample, all participants yielded a result ($t(139) = 24.994$, $p < 0.001$), and the MD was 0.83 (95% CI [0.76, 0.89]). Additionally,

all effect sizes exceeded 0.8, indicating large effects. Thus, all groups and the overall sample had significantly reached the "high" satisfaction level.

Table 8. *Levene's Test and 2-Way ANOVA for Mean Satisfaction Scores with the Service Quality of All Participants*

Levene's Test of Equality of Error Variances ^{a,b}					
	Levene Statistic	df1	df2	Sig.	
Based on Mean	.355	3	136	.786	
Based on trimmed mean	.348	3	136	.791	
2-Way ANOVA					
Descriptive Statistics-Mean		N1	N0	O1	O0
		4.35	4.31	4.44	4.22
Source	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	3	.580	4.013	.009	.081
NTCE	1	.088	.606	.438	.004
OnlineExp	1	1.607	11.125	.001	.076
NTCE * OnlineExp	1	.045	.309	.579	.002

Table 8 initially showed the null hypothesis of equal error variances could not be rejected based on Levene's test among all participants (all $p > 0.05$). Further, the 2-way ANOVA (corrected model) was significant ($p = 0.009 < 0.05$), and its $\eta_p^2 = 0.081 > 0.06$ suggested a medium effect size approaching the 0.14 of large effect size. "OnlineExp" has a significant effect ($p = 0.001 < 0.05$, $\eta_p^2 = 0.076$) indicating that preservice chemistry teachers whether or not have online learning experience had a substantial influence on their satisfaction with the service quality in the experience of this MOOC. Mean satisfaction was higher for O1 than O0 (4.44 > 4.22; $\Delta = 0.22$), but both groups were at the "high" level. By comparison, NTCE ($p = 0.438 > 0.05$) and their interaction effect ($p = 0.579 > 0.05$) did not have the significant impact. Overall, prior online learning experience was the primary driver of between-group differences in service-quality satisfaction.

Satisfaction with the Experience in the MOOC

Table 9. *One-Sample T Test and Effect Sizes for Mean Satisfaction Scores with the Experience in the MOOC*

Test Value = 3.50									
Groups	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference		Cohen's d Point Estimate	95% Confidence Interval	
					Lower	Upper		Lower	Upper
N1O1	9.130	34	< .001	.79	.61	.97	1.543	1.045	2.031
N0O1	10.180	34	< .001	.83	.66	1.00	1.721	1.191	2.241
N1O0	9.053	34	< .001	.79	.61	.97	1.530	1.034	2.016
N0O0	9.739	34	< .001	.80	.63	.96	1.646	1.129	2.152
All	19.213	139	< .001	.80	.72	.88	1.624	1.370	1.875

Table 9 demonstrated that all four groups scored significantly above the benchmark of 3.50 (all $t(34) \geq 9.053$, all $p < 0.001$), with all MDs (95% CIs) entirely above zero. For the overall sample, all participants

yielded a result ($t(139) = 19.213, p < 0.001$), and the MD was 0.80 (95% CI [0.72, 0.88]). Moreover, all effect sizes exceeded 0.8, suggesting large effects. Thus, all groups and the overall sample had significantly reached the “high” satisfaction level for the experience of this MOOC.

Table 10. *Levene’s Test and 2-Way ANOVA for Mean Satisfaction Scores with the Experience in the MOOC of All Participants*

Levene’s Test of Equality of Error Variances ^{a,b}					
	Levene Statistic	df1	df2	Sig.	
Based on Mean	.121	3	136	.948	
Based on trimmed mean	.124	3	136	.946	
2-Way ANOVA					
Descriptive Statistics-Mean		N1	N0	O1	O0
		4.29	4.32	4.31	4.30
Source	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	3	.013	.052	.984	.001
NTCE	1	.020	.082	.775	.001
OnlineExp	1	.009	.037	.849	.000
NTCE * OnlineExp	1	.009	.037	.849	.000

Table 10 initially showed the null hypothesis of equal error variances could not be rejected based on Levene’s test among all participants (all $p > 0.05$). However, the 2-way ANOVA (corrected model) was not statistically significant ($df = 3, F = 0.052, p = 0.984 > 0.05, \eta_p^2 = 0.001$). This meant that under this current survey and hypothesis testing, neither the main effects nor the interaction effect exhibited a statistically significant influence on the dependent variable. Further analysis showed that “NTCE” ($p = 0.775 > 0.05, \eta_p^2 = 0.001$), “OnlineExp” ($p = 0.849 > 0.05, \eta_p^2 \approx 0.000$) and their interaction effect ($p = 0.849 > 0.05, \eta_p^2 \approx 0.000$) all did not have a significant impact on the mean satisfaction scores with the experience. This meant that whether or not participants have passed the NTCE or whether or not they have online learning experience, neither of these two factors nor their interaction effect had a significant impact on their satisfaction with the overall experience in this MOOC.

Intention to Continue Using the MOOC

Table 11. *One-Sample T Test and Effect Sizes for Mean Scores of Intention to Continue Using the MOOC*

Test Value = 3.50									
Groups	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference		Cohen’s d Point Estimate	95% Confidence Interval	
					Lower	Upper		Lower	Upper
N1O1	15.332	34	< .001	.89	.77	1.01	2.592	1.890	3.284
N0O1	11.566	34	< .001	.86	.71	1.01	1.955	1.381	2.519
N1O0	15.705	34	< .001	.92	.80	1.04	2.655	1.940	3.360
N0O0	14.403	34	< .001	.93	.80	1.06	2.434	1.765	3.094
All	28.269	139	< .001	.90	.84	.96	2.389	2.062	2.714

Table 11 reported that all four groups scored significantly above the benchmark of 3.50 (all $t(34) \geq 11.566$, all $p < 0.001$), with all MDs (95% CIs) entirely above zero. For the overall sample, all participants yielded a result ($t(139) = 28.269$, $p < 0.001$), and the MD was 0.90 (95% CI [0.84, 0.96]). In addition, all effect sizes exceeded 0.8, suggesting large effects. Thus, all groups and the overall sample had significantly reached the “strong” intention level.

Learning Behavior in the Whole of the MOOC

Table 12. One-Sample T Test for Total Frequency of MOOC’s Videos Learning and Interactions in the Forum in the Whole of the MOOC

Total Frequency of MOOCs’ Videos Learning in the Whole of the MOOCs						
Test Value = 45						
Groups	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
N1O1	7.117	34	< .001	6.71	4.80	8.63
N0O1	10.407	34	< .001	11.14	8.97	13.32
N1O0	8.403	34	< .001	11.14	8.45	13.84
N0O0	7.161	34	< .001	8.37	6.00	10.75
All	15.990	139	< .001	9.34	8.19	10.50

Total Frequency of Interactions in the Forum in the Whole of the MOOCs						
Test Value = 30						
Groups	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
N1O1	10.195	34	< .001	18.89	15.12	22.65
N0O1	10.451	34	< .001	20.83	16.78	24.88
N1O0	9.147	34	< .001	27.11	21.09	33.14
N0O0	10.178	34	< .001	26.89	21.52	32.25
All	19.054	139	< .001	23.43	21.00	25.86

Table 12 demonstrated two-part results of the learning behavior. First, for the MOOC’s videos learning, all four groups scored significantly above the benchmark of 45 (all $t(34) \geq 7.117$, all $p < 0.001$), with all MDs (95% CIs) entirely above zero. All participants yielded a result ($t(139) = 15.990$, $p < 0.001$), and the MD was 9.34 (95% CI [8.19, 10.50]). Hence, all participants significantly met the required standard for total video-learning frequency. Next, for the interactions in the forum, all four groups also scored significantly above the benchmark of 30 (all $t(34) \geq 9.147$, all $p < 0.001$), with all MDs (95% CIs) entirely above zero. All participants yielded a result ($t(139) = 19.054$, $p < 0.001$), and the MD was 23.43 (95% CI [21.00, 25.86]). Therefore, they also met the required standard for the total frequency of interactions.

Exam Results in the Whole of the MOOC

Table 13. One-Sample T Test for the "Comprehensive Final Scores" in the Whole of the MOOC

Groups	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
					Test Value = 85.00	
N1O1	4.664	34	< .001	4.42	2.49	6.34
N0O1	4.118	34	< .001	3.69	1.87	5.51
N1O0	4.164	34	< .001	3.87	1.98	5.75
N0O0	3.839	34	< .001	3.63	1.71	5.56
All	8.471	139	< .001	3.90	2.99	4.81

Table 13 revealed that all four groups scored significantly above the benchmark of 85.00 (all $t(34) \geq 3.839$, all $p < 0.001$), with all MDs (95% CIs) entirely above zero. For the overall sample, all participants yielded a result ($t(139) = 8.471$, $p < 0.001$), and the MD was 3.90 (95% CI [2.99, 4.81]). Accordingly, all groups and the overall participants met the required "passing" standard and even reached the "excellent" level.

Satisfaction with the Overall Benefits & Overall Satisfaction with the Whole of the MOOC

Table 14. One-Sample T Test and Effect Sizes for Mean Satisfaction Scores with the Overall Benefits across the Entire Learning Process in the MOOC

Groups	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference		Cohen's d Point Estimate	95% Confidence Interval	
					Lower	Upper		Lower	Upper
					Test Value = 4.50				
N1O1	8.921	34	< .001	.25	.20	.31	1.508	1.016	1.990
N0O1	8.005	34	< .001	.25	.19	.31	1.353	.887	1.809
N1O0	8.243	34	< .001	.26	.20	.32	1.393	.920	1.855
N0O0	7.503	34	< .001	.25	.19	.32	1.268	.816	1.710
All	16.419	139	< .001	.25	.22	.28	1.388	1.154	1.619

Table 14 revealed that all four groups scored significantly above the benchmark of 4.50 (all $t(34) \geq 7.503$, all $p < 0.001$), with all MDs (95% CIs) entirely above zero. For the overall sample, all participants yielded a result ($t(139) = 16.419$, $p < 0.001$), and the MD was 0.25 (95% CI [0.22, 0.28]). Thus, all groups and the overall sample had significantly reached the "very high" satisfaction level for the overall benefits in this MOOC. Moreover, all large effect sizes suggested the fact that various good and high-quality aspects of the whole of this MOOC could endow learners with extremely positive gains.

Table 15. One-Sample T Test and Effect Sizes for Mean Overall Satisfaction Scores with the Whole of the MOOC

Test Value = 4.50									
Groups	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference		Cohen's d Point Estimate	95% Confidence Interval	
					Lower	Upper		Lower	Upper
N1O1	8.863	34	< .001	.26	.20	.32	1.498	1.008	1.978
N0O1	7.740	34	< .001	.25	.18	.31	1.308	.849	1.757
N1O0	7.985	34	< .001	.25	.19	.32	1.350	.884	1.805
N0O0	7.516	34	< .001	.26	.19	.33	1.270	.818	1.713
All	16.139	139	< .001	.26	.22	.29	1.364	1.132	1.593

Similarly, according to Table 15, all four groups and the overall sample scored significantly above the benchmark of 4.50 (all $p < 0.001$, all MDs (95% CIs) > 0), revealing that they also reached the “very high” satisfaction level for the overall MOOC. All large effect sizes suggested that various good and high-quality aspects of the whole of this MOOC could provide learners with a positive overall learning process.

Table 16. Levene's Test and 2-Way ANOVA for Mean Satisfaction Scores with the Overall Benefits across the Entire Learning Process in the MOOC of All Participants

Levene's Test of Equality of Error Variances ^{a,b}				
	Levene Statistic	df1	df2	Sig.
Based on Mean	.200	3	136	.896
Based on trimmed mean	.169	3	136	.917

2-Way ANOVA				
Descriptive Statistics-Mean	N1	N0	O1	O0
	4.76	4.75	4.75	4.76

Source	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	3	.001	.022	.995	.000
NTCE	1	.001	.033	.855	.000
OnlineExp	1	.001	.033	.855	.000
NTCE * OnlineExp	1	.000	.000	1.000	.000

Table 16 firstly showed the result of equal error variances by Levene's test among all participants (all $p > 0.05$). Nevertheless, the 2-way ANOVA (corrected model) was not statistically significant ($df = 3$, $F = 0.022$, $p = 0.995 > 0.05$, $\eta_p^2 \approx 0.000$). Further, “NTCE” and “OnlineExp” (both $p = 0.855 > 0.05$, $\eta_p^2 = 0.001$), and their interaction effect ($p = 1.000 > 0.05$, $\eta_p^2 \approx 0.000$) all did not have a significant impact. This meant that whether or not participants have passed the NTCE or whether or not they have online learning experience, neither of these two factors nor their interaction effect had a significant impact on their satisfaction with the overall benefits across the entire learning process in the MOOC.

Table 17. *Levene’s Test and 2-Way ANOVA for Mean Overall Satisfaction Scores with the Whole of the MOOC of All Participants*

Levene’s Test of Equality of Error Variances ^{a,b}					
	Levene Statistic	df1	df2	Sig.	
Based on Mean	.447	3	136	.720	
Based on trimmed mean	.446	3	136	.720	
2-Way ANOVA					
Descriptive Statistics-Mean	N1	N0	O1	O0	
	4.76	4.75	4.75	4.76	
Source	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	3	.001	.029	.993	.001
NTCE	1	.000	.008	.929	.000
OnlineExp	1	.000	.008	.929	.000
NTCE * OnlineExp	1	.003	.072	.789	.001

Likewise, in terms of Table 17, error variances were equal (all $p > 0.05$). Nonetheless, corrected model was not significant ($p = 0.993 > 0.05$, $\eta_p^2 = 0.001$). Whether or not participants have passed the NTCE or whether or not they have online learning experience (both $p = 0.929 > 0.05$, $\eta_p^2 \approx 0.000$), neither of these two factors nor their interaction effect ($p = 0.789 > 0.05$, $\eta_p^2 = 0.001$) had a significant impact on their overall satisfaction with the whole of the MOOC.

Intention to Review the MOOC

Table 18. *One-Sample T Test and Effect Sizes for Mean Scores of Intention to Review the MOOC*

Test Value = 3.50									
Groups	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference		Cohen’s d Point Estimate	95% Confidence Interval	
					Lower	Upper		Lower	Upper
N1O1	20.156	34	< .001	1.04	.94	1.15	3.407	2.531	4.274
N0O1	19.049	34	< .001	1.02	.91	1.13	3.220	2.385	4.046
N1O0	19.819	34	< .001	1.06	.95	1.17	3.350	2.487	4.205
N0O0	18.092	34	< .001	1.03	.91	1.14	3.058	2.258	3.849
All	38.868	139	< .001	1.04	.98	1.09	3.285	2.864	3.704

Lastly, Table 18 showed that all four groups and the overall sample scored significantly above the benchmark of 3.50 (all $p < 0.001$, all MDs (95% CIs) > 0), and that all effect sizes were large, revealing that they all reached the “strong” intention level for reviewing this MOOC. Besides, the observed means were closely near 4.50 (threshold for “very strong”) but did not achieve statistical significance.

DISCUSSION

Following the satisfaction-focused updated ISSM and the joint analytical approach, the comprehensive findings suggested that: preservice chemistry teachers showed at least “high” satisfaction across all aspects of this Instructional Design for Secondary School Chemistry MOOC; and, as the course progressed, their satisfaction with the overall benefits and their overall satisfaction with the whole of

this MOOC ultimately reached a “very high” level. This progression was both natural and logical. It represented a shift from short-term phase-level approval to comprehensive recognition and reflected deeper understanding and identification developed over this MOOC in the long-term learning process. This finding aligns with Hsu’s (2021) ten-week multilevel modeling study of MOOC learners, which showed that time significantly shapes engagement; as the course progressed, engagement rose, thereby enhancing satisfaction (Yang & Jiang, 2020). These upward trajectories suggested growing familiarity, deeper understanding, and stronger identification (Hsu, 2021; Yang & Jiang, 2020). When learners pursued personally interesting content and meaningful interactions, satisfaction tended to increase with sustained engagement over time (Bahati et al., 2019). Additionally, learning behavior, exam results, and intentions measured at two time points likewise offered strong support for these satisfaction findings.

In further detail, preservice chemistry teachers with prior online learning experience reported significantly higher satisfaction with system and service quality than those without, whereas NTCE-passed status showed no significant differences. From the perspective of accumulated experience in online learning, preservice teachers with prior experience may be more familiar with and better understand how to use MOOCs. They are also likely to hold higher expectations for aspects of system quality such as platform stability, usability, and technical support; when met, they may rate system quality more positively (Alshammari & Alshammari, 2024; Elkins et al., 2021). Likewise, in terms of sensitivity to service quality, those with extensive online learning experience might be more attuned to and better able to perceive and rate the quality of services; when the MOOC’s services are excellent, this could lead to higher service-quality satisfaction as well (Taat & Umbit, 2021). Thus, despite generally high satisfaction, preservice chemistry teachers with online learning experience further evidenced excellent system stability, user-friendliness, interactive responsiveness, and developer support, validating this MOOC’s well-executed, user-centered design. Nevertheless, NTCE-passed status primarily reflects professional competencies in areas such as educational theory, instructional skills, and classroom management, rather than familiarity with technology use or ability to assess service experiences (Yin, 2023).

For information quality, conversely, preservice chemistry teachers who had passed the NTCE reported significantly higher satisfaction compared to those who had not, whereas online learning experience did not show such differences. From the perspective of the teacher certification being an important credential for validating professional competence, those who have passed tend to possess more evaluative experience and a stronger sense of professional identity (Gu & Meng, 2024). Consequently, they may hold higher expectations and standards regarding content rigor and structure, information accuracy, and course materials; when met, they are likely to express greater satisfaction with information quality (Angelova, 2021). Also, certified preservice teachers have undergone more systematic teacher training and examination, enhancing their ability to more likely to identify and judge the quality of educational information (Luo & Lu, 2020). Hence, despite generally high satisfaction, those who passed NTCE further evidenced excellent disciplinary rigour in chemistry instructional design, scientific soundness in pedagogical guidance, and accuracy and practical relevance in its information. Conversely, online learning experience primarily influences users’ operational use and perception of the platform (Du et al., 2021).

Moreover, all their interaction effects were not significant indicating that the impacts of NTCE-passed status and online learning experience on satisfaction were independent, evidencing that this MOOC was not particularly advantageous to any specific combination of learner’s background. Notably, it achieved higher recognition from more professionally trained preservice chemistry teachers while effectively meeting the learning needs of those from different backgrounds.

In comparison, across NTCE-passed status, online learning experience, and their interaction, there were no significant differences in their satisfaction with the experience, satisfaction with the overall benefits, or overall satisfaction with the whole of the MOOC. These findings, unlike the results related to satisfaction with qualities, were justifiable. Subjective experiential and overall satisfaction were shaped more by the overall adaptability and inclusiveness of the MOOC design, and also the actually practical

perception, thereby reducing the influence of background differences (Zamri et al., 2021; Ismailov & Chiu, 2022). Meanwhile, given the generally high satisfaction, a ceiling effect might have compressed between-group differences.

The contribution of this study lay not only in reporting satisfaction across this MOOC's dimensions, but also in offering a reusable and testable lens for benchmarking and between-group analysis in online-course satisfaction research. In recent years, international research has also developed comparable, satisfaction-centered frameworks for assessing online-course satisfaction. From a service-quality perspective, the SERVQUAL dimensions were used to characterize and predict e-learning learners' satisfaction as a core outcome (Sumi & Kabir, 2021), while, within the Community of Inquiry framework, satisfaction was positioned as a key endpoint and the pathway across teaching, social, and cognitive presence was articulated to explain how satisfaction emerged and increased (Farrokhnia et al., 2025). At the analytical level, recent research has likewise provided additional multi-timescale points of reference. Some adopted a two-wave longitudinal design and used cross-lagged structural equation modeling to test lagged effects on satisfaction, thereby capturing the dynamics of satisfaction, as illustrated by Besalti and Satici (2022); others adopted a stage-based approach within one online course, collecting learners' coded evaluative comments over several weeks and applying MANOVA to compare stage differences, as demonstrated by Wei et al. (2022). Moreover, beyond MOOCs as a representative type of online courses, the workflow proposed in this study could be extended to contexts such as synchronous online classes and small private online courses (El Moussaddar et al., 2025; Hamoen et al., 2022), thereby facilitating cross-format validation.

CONCLUSION

From this quantitative practical case, it can be seen that this satisfaction-focused ISSM can be effectively used for long-term, longitudinal, multi-stage explorations of satisfaction. The joint analytical approach effectively achieves the evaluation objective of first "benchmarking to determine whether the standard is met", and then "revealing between-group differences to judge the adaptability of the online course". At the same time, the integrated findings indicated that this MOOC is suitable for broad adoption and holds considerable practical value and promising prospect for wider application in enhancing preservice chemistry teachers' instructional design abilities and improving their course satisfaction through the technologically supported training mode. For future research, it is recommended to use this framework with additional cases to evaluate online-course satisfaction, especially for courses oriented toward specific contextual groups as supplement. Moreover, when learners present more background factors, reasonably consider replacing 2-way ANOVA with multi-factor ANOVA or ANCOVA.

REFERENCES

- Alfaki, I. A. (2021). Delone and McLean information systems success model in a blended-learning context. *International Journal of Information and Communication Technology Education (IJICTE)*, 17(4), 1–17. <https://doi.org/10.4018/IJICTE.20211001.0a18>
- Alkharusi, H. (2022). A descriptive analysis and interpretation of data from likert scales in educational and psychological research. *Indian Journal of Psychology and Education*, 12(2), 13–16. https://sqa.elsevierpure.com/ws/portalfiles/portal/62602494/A_107_2022.pdf
- Alshammari, S. H., & Alshammari, R. A. (2024). An integration of expectation confirmation model and information systems success model to explore the factors affecting the continuous intention to utilise virtual classrooms. *Scientific Reports*, 14(1), 18491. <https://doi.org/10.1038/s41598-024-69401-8>
- Andrade, C. (2024). Understanding factorial designs, main effects, and interaction effects: Simply explained with a worked example. *Indian Journal of Psychological Medicine*, 46(2), 175–177. <https://doi.org/10.1177/02537176241237066>
- Angelova, S. (2021). Development of student teachers' satisfaction with e-learning science courses. In *AIP Conference Proceedings* (Vol. 2333, No. 1, p. 050011). AIP Publishing LLC. <https://doi.org/10.1063/5.0042071>
- Anić, P., & Tončić, M. (2024). Affect and satisfaction time dependency: An experience sampling

- study. *PsyCh Journal*, 13(3), 456–464. <https://doi.org/10.1002/pchj.763>
- Bahati, B., Fors, U., Hansen, P., Nouri, J., & Mukama, E. (2019). Measuring learner satisfaction with formative e-assessment strategies. *International Journal of Emerging Technologies in Learning (IJET)*, 14(07), 61–79. <https://doi.org/10.3991/ijet.v14i07.9120>
- Berlin, K., & Weaver, K. V. (2022). Teaching strategies students find helpful in online learning courses. *College Teaching*, 70(3), 319–327. <https://doi.org/10.1080/87567555.2021.1940814>
- Besalti, M., & Satici, S. A. (2022). Online learning satisfaction and internet addiction during COVID-19 pandemic: A two-wave longitudinal study. *TechTrends*, 66(5), 876–882. <https://doi.org/10.1007/s11528-022-00697-x>
- Cheng, X., Bao, Y., Yang, B., Chen, S., Zuo, Y., & Siponen, M. (2023). Investigating students' satisfaction with online collaborative learning during the COVID-19 period: An expectation-confirmation model. *Group Decision and Negotiation*, 32(4), 749–778. <https://doi.org/10.1007/s10726-023-09829-x>
- Chiu, K. C., & Chen, R. C. (2024). The analysis of learning performance satisfaction for physical and online learning: A case study from Taiwan. *Electronic Journal of E-Learning*, 22(9), 28–45. <https://doi.org/10.34190/ejel.22.9.3305>
- Dai, J., Zhang, X., & Wang, C. (2024). A meta-analysis of learners' continuance intention toward online education platforms. *Education and Information Technologies*, 29(16), 21833–21868. <https://doi.org/10.1007/s10639-024-12654-7>
- DeLone, W. H., & McLean, E. R. (1992). Information systems success: The quest for the dependent variable. *Information Systems Research*, 3(1), 60–95. <https://doi.org/10.1287/isre.3.1.60>
- DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success: A ten-year update. *Journal of Management Information Systems*, 19(4), 9–30. <https://doi.org/10.1080/07421222.2003.11045748>
- Dong, N., Kelcey, B., & Spybrook, J. (2023). Experimental design and power for moderation in multisite cluster randomized trials. *The Journal of Experimental Education*, 92(4), 741–757. <https://doi.org/10.1080/00220973.2023.2226934>
- Du, Z., Wang, F., & Wang, S. (2021). Reviewer experience vs. expertise: Which matters more for good course reviews in online learning? *Sustainability*, 13 (21), 12230. <https://doi.org/10.3390/su132112230>
- El Moussaddar, J., Belfakir, L., Tadist, K., Habiballah, S., & Bellet, M. (2025). The impact of social interaction on students' satisfaction with synchronous online learning at FLDM. *Social Sciences & Humanities Open*, 12, 101759. <https://doi.org/10.1016/j.ssaho.2025.101759>
- Elkins, R., Mcdade, R., Collins, M., & King, M. (2021). Student satisfaction with online learning experiences associated with COVID-19. *Research Directs in Health Sciences*, 1(1), 9. <http://dx.doi.org/10.53520/rdhs2021.10419>
- Elmunyah, H., Nafalski, A., Wibawa, A. P., & Dwiyanto, F. A. (2023). Understanding the impact of a learning management system using a novel modified DeLone and McLean model. *Education Sciences*, 13(3), 235. <https://doi.org/10.3390/educsci13030235>
- Farrokhnia, M., Taghizade, A., Ahmadi, R., Papadopoulos, P. M., & Noroozi, O. (2025). Community of inquiry: A bridge linking motivation and self-regulation to satisfaction with e-learning. *The Internet and Higher Education*, 65, 100992. <https://doi.org/10.1016/j.iheduc.2025.100992>
- Goss-Sampson, M. A. (2022). *Statistical analysis in JASP: A guide for students* (5th ed., version 0.16.1). JASP. <https://jasp-stats.org/wp-content/uploads/2022/04/Statistical-Analysis-in-JASP-A-Students-Guide-v16.pdf>
- Gu, C., & Meng, Q. (2024). Research on the professional identity and improvement strategies of undergraduate students in primary education major. *Journal of Liaoning Normal University (Social Science Edition)*, 47(1), 90–97. <https://doi.org/10.16216/j.cnki.lsxwbk.202401090>
- Hamad, F., Shehata, A., & Al Hosni, N. (2024). Predictors of blended learning adoption in higher education institutions in Oman: Theory of planned behavior. *International Journal of Educational Technology in Higher Education*, 21(13), 1–28. <https://doi.org/10.1186/s41239-024-00443-8>
- Hamoen, E. C., De Jong, P. G., Van Blankenstein, F. M., & Reinders, M. E. (2022). Design and first impressions of a small private online course in clinical workplace learning: Questionnaire and interview study. *JMIR Medical Education*, 8(2), e29624. <https://doi.org/10.2196/29624>
- Hsu, L. (2021). Learning tourism and hospitality subjects with massive open online courses (MOOCs):

- A cross-sectional and longitudinal study. *Journal of Hospitality, Leisure, Sport & Tourism Education*, 29, 100276. <https://doi.org/10.1016/j.jhlste.2020.100276>
- Ismailov, M., & Chiu, T. K. (2022). Catering to inclusion and diversity with universal design for learning in asynchronous online education: A self-determination theory perspective. *Frontiers in Psychology*, 13, 819884. <https://doi.org/10.3389/fpsyg.2022.819884>
- Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. *Educational and Psychological Measurement*, 30(3), 607–610. <https://doi.org/10.1177/001316447003000308>
- Lai, C., & Jin, T. (2021). Teacher professional identity and the nature of technology integration. *Computers & Education*, 175, 104314. <https://doi.org/10.1016/j.compedu.2021.104314>
- Li, L., Wang, Q., & Li, J. (2022). Examining continuance intention of online learning during COVID-19 pandemic: Incorporating the theory of planned behavior into the expectation–confirmation model. *Frontiers in Psychology*, 13, 1046407. <https://doi.org/10.3389/fpsyg.2022.1046407>
- Liu, Y. (2024). *Instructional design for secondary school chemistry* [MOOC]. Chinese University MOOC. <https://www.icourse163.org/course/HENANNU-1001796020>
- Lu, H., & Smiles, R. (2024). Redesign of an online course for student success: A case study. *Quest Journals: Journal of Research in Humanities and Social Science*, 12(4), 78–84. <https://www.questjournals.org/jrhss/papers/vol12-issue4/12040107.pdf>
- Luik, P., & Taimalu, M. (2021). Predicting the intention to use technology in education among student teachers: A path analysis. *Education Sciences*, 11(9), 564. <https://doi.org/10.3390/educsci11090564>
- Luo, X. L., & Lu, Y. H. (2020). A study on the cultivation of cultural literacy among normal university students under the new teacher certification examination context. *Journal of Shandong Institute of Commerce and Technology*, (5), 54–56, 69. <https://doi.org/10.13396/j.cnki.jsict.2020.05.013>
- Martin, F., & Bolliger, D. U. (2022). Developing an online learner satisfaction framework in higher education through a systematic review of research. *International Journal of Educational Technology in Higher Education*, 19(1), 50. <https://doi.org/10.1186/s41239-022-00355-5>
- Mathew, P. V., Pappachan, J., & Kushwaha, S. (2024). The interplay of online attributes and learning modes in shaping satisfaction and effectiveness in online learning. *Discover Education*, 3(1), 241. <https://doi.org/10.1007/s44217-024-00344-y>
- Ministry of Education of the People's Republic of China. (2020). *Chemistry curriculum standards for general high school* (2017 edition, 2020 revision). https://www.htu.edu.cn/hxxy/_t251/2021/1025/c17076a222541/page.htm
- Ministry of Education of the People's Republic of China. (2021). *Professional competency standards for secondary education preservice teachers* (trial). www.moe.gov.cn/srcsite/A10/s6991/202104/t20210412_525943.html
- Ministry of Education of the People's Republic of China. (2022). *Chemistry curriculum standards for compulsory education* (2022 edition). <https://www.pep.com.cn/ebook/2022yjkcbz/hx/mobile/index.html>
- Mustafa, A. S., & Ali, N. (2023). The adoption and use of Moodle in online learning: A systematic review. *Information Sciences Letters*, 12(1), 341–351. <http://dx.doi.org/10.18576/isl/120129>
- Pham, D. M., Vanacore, K. P., Sales, A. C., & Gagnon-Bartsch, J. A. (2024). LOOL: Towards personalization with flexible and robust estimation of heterogeneous treatment effects. In B. Paaßen & C. Demmans Epp (Eds.), *Proceedings of the 17th International Conference on Educational Data Mining* (pp. 376–384). International Educational Data Mining Society. <https://doi.org/10.5281/zenodo.12729840>
- Ponce-Renova, H. F. (2022). Comparing effect sizes and their confidence intervals: A primer on equivalence testing in educational research. *Journal of New Approaches in Educational Research*, 11(2), 209–225. <https://doi.org/10.7821/naer.2022.7.930>
- Rahmani, A. M., Groot, W., & Rahmani, H. (2024). Dropout in online higher education: A systematic literature review. *International Journal of Educational Technology in Higher Education*, 21(1), 19. <https://doi.org/10.1186/s41239-024-00450-9>
- Reichgelt, J., & Smith, T. C. (2024). The effect of course delivery mode on student performance and student satisfaction: A case study. *Trends in Higher Education*, 3(4), 872–884. <https://doi.org/10.3390/higheredu3040050>

- Sabeh, H. N., Husin, M. H., Kee, D. M. H., Baharudin, A. S., & Abdullah, R. (2021). A systematic review of the DeLone and McLean model of information systems success in an e-learning context (2010–2020). *IEEE Access*, *9*, 81210–81235. <https://doi.org/10.1109/ACCESS.2021.3084815>
- Schriek, J., Carstensen, B., Soellner, R., & Klusmann, U. (2024). Pandemic rollercoaster: University students' trajectories of emotional exhaustion, satisfaction, enthusiasm, and dropout intentions pre-, during, and post-COVID-19. *Teaching and Teacher Education*, *148*, 104709. <https://doi.org/10.1016/j.tate.2024.104709>
- Sumi, R. S., & Kabir, G. (2021). Satisfaction of e-learners with electronic learning service quality using the SERVQUAL model. *Journal of Open Innovation: Technology, Market, and Complexity*, *7*(4), 227. <https://doi.org/10.3390/joitmc7040227>
- Sureiman, O., & Mangera, C. M. (2020). Conceptual framework on conducting two-way analysis of variance. *Journal of the Practice of Cardiovascular Sciences*, *6*(3), 207–215. https://doi.org/10.4103/jpcs.jpcs_75_20
- Taat, M. S., & Umbit, A. F. (2021). E-learning: satisfaction, acceptance and experience among pre-service teachers in Malaysia. *Advanced Journal of Technical and Vocational Education*, *5*(4), 01–07. <https://doi.org/10.26666/rmp.ajtve.2021.4.1>
- Tian, M., & Lu, G. (2022). Online learning satisfaction and its associated factors among international students in China. *Frontiers in Psychology*, *13*, 916449. <https://doi.org/10.3389/fpsyg.2022.916449>
- Tipton, E., & Olsen, R. B. (2022). *Enhancing the generalizability of impact studies in education*. Mathematica Policy Research. <https://econpapers.repec.org/RePEc:mpr:mprres:35d5625333dc480aba9765b3b683af2f>
- Wei, W., Yu, Y., & Gao, G. (2022). Investigating learners' changing expectations on learning experience in a MOOC of professional translation and interpreter training. *Sage Open*, *12*(4), Article 21582440221134577. <https://doi.org/10.1177/21582440221134577>
- Yang, D. S., & Jiang, K. (2020). Research on the influencing factors of the continuance intention online education platforms: based on expectation confirmation theory. *Journal of Management and Humanity Research*, *3*, 61–70. <https://doi.org/10.22457/jmhr.v03a06106>
- Yin, X., Zhang, J., Li, G., & Luo, H. (2024). Understanding learner satisfaction in virtual learning environments: Serial mediation effects of cognitive and social-emotional factors. *Electronics*, *13*(12), 2277. <https://doi.org/10.3390/electronics13122277>
- Yin, Y. L. (2023). An analysis of strategies for teacher certification examination among vocational normal students under the context of the "National Examination". *Shanxi Youth*, (20), 49–51. <https://doi.org/CNKI:SUN: SXQS.0.2023-20-016>
- Yu, Q. (2022). Factors influencing online learning satisfaction. *Frontiers in Psychology*, *13*, 852360. <https://doi.org/10.3389/fpsyg.2022.852360>
- Zamri, N., Omar, N. B., Anwar, I. S. K., & Fatzel, F. H. M. (2021). Factors affecting students' satisfaction and academic performance in open & distance learning (ODL). *International Journal of Academic Research in Business and Social Sciences*, *11*(11), 1–16. <http://dx.doi.org/10.6007/IJARBS/v11-i11/11194>
- Zeng, X., & Wang, T. (2021). College student satisfaction with online learning during COVID-19: A review and implications. *International Journal of Multidisciplinary Perspectives in Higher Education*, *6*(1), 182–195. <https://doi.org/10.32674/jimphe.v6i1.3502>
- Zhang, M., Li, S., & Zhang, Y. (2023). A meta-analysis of the moderating role of prior learning experience and mandatory participation on factors influencing MOOC learners' continuance intention. *Australasian Journal of Educational Technology*, *39*(2), 115–141. <https://doi.org/10.14742/ajet.7795>
- Zhou, L. (2024). A study on the factors influencing online learning satisfaction based on the customer satisfaction model and the 3P learning model—The mediating effect of mind-flow experience. *Statistics and Applications*, *13*(5), 2032–2043. <https://doi.org/10.12677/sa.2024.135197>
- Zhu, M., Sari, A. R., & Lee, M. M. (2022). Trends and issues in MOOC learning analytics empirical research: A systematic literature review (2011–2021). *Education and Information Technologies*, *27*(7), 10135–10160. <https://doi.org/10.1007/s10639-022-11031-6>

Zygmunt, C. S. (2023). Managing the assumption of normality within the general linear model with small samples: Guidelines for researchers regarding if, when and how. *The Quantitative Methods for Psychology*, 19(4), 302–332. <https://doi.org/10.20982/tqmp.19.4.p302>