

Multimodal Cognitive-state Sensing for Cognitive Load Assessment in Preservice Elementary Teachers' Instructional Design Tasks

Qin Lai,¹ Hui-Chen Tsai,² Shiyi Wang,¹ and Cheng-Fu Yang^{3,4*}

¹School of Language and Culture, Dongguan City University, Guangdong Province 523419, China

²Center for General Education, Taiwan Steel University of Science and Technology, Kaohsiung 821, Taiwan

³Department of Chemical and Materials Engineering, National University of Kaohsiung, Kaohsiung 811, Taiwan

⁴Department of Aeronautical Engineering, Chaoyang University of Technology, Taichung 413, Taiwan

(Received July 9, 2026; accepted September 1, 2026)

Keywords: multimodal sensing, physiological sensing, eye-tracking sensor, cognitive load monitoring, sensor fusion, smart teacher education

Instructional design is a cognitively demanding activity in preservice teacher education because it requires the integration of instructional objectives, learning activities, assessment strategies, and pedagogical decision-making. Conventional cognitive load studies have mainly relied on post-task self-report instruments, which are insufficient for capturing real-time cognitive fluctuations during task execution. In this study, we proposed a sensor-integrated multimodal cognitive load assessment framework for instructional design tasks in preservice elementary teachers by combining physiological sensing signals, eye-tracking sensors, behavioral interaction sensing, synchronized multimodal data fusion, subjective cognitive load scales, and expert evaluations of instructional design quality. The study included 72 participants and three instructional design subtasks: instructional objective formulation, learning activity planning, and learning assessment design. A real-time multimodal sensing architecture was established to support synchronized signal acquisition, cross-modal feature extraction, and dynamic cognitive load monitoring during instructional design processes. The results revealed significant differences among the three subtasks across all four subjective cognitive load dimensions. Learning assessment design consistently produced the highest cognitive load, followed by instructional objective formulation, whereas learning activity planning showed the lowest load level. These findings demonstrate the potential of multimodal educational sensing technologies and sensor-fusion-based cognitive monitoring for task-sensitive cognitive load assessment and smart teacher education applications. The framework also provides application requirements for future educational sensors, including wearable sensing materials and compact multimodal integration. However, the single-institution sample and limited correspondence between subjective and objective sensing indicators constrain generalizability. Future studies should improve sensor-fusion methods and validate the framework with larger and more diverse samples.

*Corresponding author: e-mail: cfyang@nuk.edu.tw
<https://doi.org/10.18494/SAM6516>

1. Introduction

Instructional design is a fundamental competency in preservice teacher education because it requires learners to formulate instructional objectives, organize learning activities, and align assessment strategies with intended learning outcomes.^(1–4) Unlike simple writing or planning activities, instructional design involves continuous pedagogical reasoning, information integration, and instructional decision-making. Preservice elementary teachers must simultaneously consider curriculum standards, learner characteristics, instructional strategies, classroom interactions, and assessment coherence while constructing teaching plans. Such multidimensional processing imposes substantial cognitive demands because learners repeatedly evaluate, revise, and coordinate multiple instructional elements during task execution. Among these components, learning assessment design is particularly demanding because it requires preservice teachers to revisit instructional objectives and learning activities while generating appropriate evidence for evaluating student learning outcomes. Consequently, instructional design can be regarded as a dynamic cognitive activity characterized by continuous mental workload adaptation and complex decision-making behaviors.

With the rapid advancement of smart education technologies and sensor-assisted learning environments, cognitive load assessment has gradually evolved from conventional questionnaire-based evaluation toward intelligent educational sensing systems. Educational sensors and multimodal sensing analytics have become increasingly important in learning analytics, adaptive instruction, and human–computer interaction research because they enable the continuous monitoring of learner cognitive states during task execution. From the perspective of smart sensing systems, instructional design activities generate diverse physiological, visual attention, and behavioral interaction signals that provide valuable information regarding cognitive processing and pedagogical reasoning. Recent developments in wearable sensing devices, eye-tracking technologies, and behavioral interaction monitoring platforms have further encouraged the development of sensor-based cognitive monitoring approaches capable of supporting adaptive educational environments and intelligent teacher-training systems.

Self-report cognitive load scales remain widely used because they are easy to administer and directly reflect learners' perceived effort and task difficulty.^(5–7) However, instructional design is inherently a multi-stage, process-sensitive, and temporally dynamic activity. A single post-task questionnaire cannot fully capture transient hesitation, repeated revision behavior, visual search transitions, or physiological fluctuations occurring during task execution. Learners may experience rapid cognitive changes while refining instructional objectives, selecting learning activities, or constructing assessments, yet these temporary variations are often overlooked when relying solely on retrospective self-report measurements. Furthermore, subjective ratings may be affected by interpretation differences, memory bias, and self-perception variability, limiting their sensitivity in detecting subtle cognitive state transitions. Therefore, traditional self-report approaches alone are insufficient for constructing comprehensive and continuous cognitive load evaluation models for instructional design tasks.

To overcome these limitations, multimodal sensing has emerged as an effective process-oriented approach for cognitive load assessment by integrating physiological sensing technologies, eye-tracking sensors, and behavioral process analytics.^(8–17) Within multimodal

educational sensing systems, physiological sensing signals can reflect arousal levels, stress responses, and mental workload variations associated with cognitive processing. Eye-tracking sensors provide information regarding visual attention allocation, fixation duration, gaze transitions, and mental effort during instructional planning activities. In addition, behavioral interaction sensing logs reveal hesitation patterns, interaction efficiency, revision frequency, and temporal task progression. Compared with conventional self-report approaches, multimodal sensing frameworks offer more objective, continuous, and temporally sensitive evidence for differentiating instructional task demands and identifying learner cognitive response patterns. The integration of heterogeneous sensing signals through synchronized multimodal sensor fusion further enables the exploration of hidden learner characteristics and adaptive learning behaviors that may not be observable through traditional assessment methods alone. The sensing modalities adopted in this study were not developed entirely from scratch but were established on previous physiological, eye-tracking, and multimodal cognitive-load sensing approaches.^(8–14,16) The present study extends these approaches by synchronizing physiological, eye-tracking, and behavioral sensing streams within a unified framework specifically for cognitive-state assessment during preservice teachers' instructional design tasks.

Recent studies have increasingly applied multimodal sensing technologies in learning analytics, adaptive tutoring systems, and educational human–computer interaction environments. Nevertheless, limited research has specifically examined multimodal sensor-based cognitive load assessment in preservice teacher instructional design contexts. Existing investigations have mainly focused on general learning tasks, online learning engagement, or problem-solving environments, whereas fewer studies have explored how different instructional design subtasks generate distinct cognitive sensing patterns and temporal sensing responses. Moreover, the relationships between subjective cognitive perceptions and objective sensor-derived indicators remain insufficiently understood in instructional planning activities. Such investigations are important because cognitive load variations may directly affect instructional design quality, pedagogical reasoning performance, and teaching preparation outcomes. From the perspective of smart teacher education, sensor-assisted cognitive monitoring systems may contribute to adaptive instructional support platforms capable of identifying learners who require additional scaffolding or personalized guidance during instructional planning tasks.

Accordingly, we proposed a sensor-integrated multimodal cognitive load assessment framework for instructional design tasks in preservice elementary teachers. The proposed framework integrated physiological sensing signals, eye-tracking indicators, behavioral interaction sensing logs, subjective cognitive load scales, and expert evaluations of instructional design quality to establish a comprehensive educational sensing model for cognitive assessment. In addition, a real-time multimodal sensing architecture was developed to support synchronized signal acquisition, cross-modal feature extraction, multimodal sensor fusion, and dynamic cognitive load monitoring during instructional design processes. Specifically, we investigated whether different instructional design subtasks generated distinct cognitive load characteristics, whether subjective ratings corresponded with objective sensing indicators, whether practicum experience was associated with lower cognitive load and higher instructional design quality, and whether multimodal sensing features can support exploratory learner profiling. The primary contribution of this study lies in extending instructional design cognitive load assessment from a single-source self-report approach toward a multimodal educational sensing framework

characterized by continuous monitoring, synchronized sensor fusion, process sensitivity, and intelligent sensing analytics capability. Furthermore, the proposed framework demonstrates the feasibility of integrating educational sensors and multimodal cognitive analytics into smart teacher education systems for future adaptive learning, intelligent cognitive monitoring, and sensor-assisted instructional support applications. Accordingly, the present study focuses on multimodal cognitive-state sensing and computational analysis rather than mediation modeling, with particular emphasis on task-sensitive cognitive load assessment and learner cognitive-state profiling.

2. Methodology

The study included 72 preservice elementary teachers recruited from an elementary teacher education program at a university through convenience sampling. Participation was voluntary. The participants included 52 females and 20 males, with a mean age of 24.5 years ($SD = 2.3$; range = 22–31 years). Most participants had completed their undergraduate education and were enrolled in preservice elementary teacher education or teacher-training programs. Therefore, subject specialization was not used as a primary grouping variable in this study. Among the participants, 36 had practicum experience and 36 had no practicum experience. The overall experimental task was instructional design, which consisted of three sequential subtasks: instructional objective formulation, learning activity planning, and learning assessment design. In the first subtask, participants formulated instructional objectives according to the given teaching topic and intended learning outcomes. In the second subtask, they planned learning activities and instructional procedures aligned with these objectives. In the third subtask, they designed assessment methods and criteria to evaluate whether the intended learning outcomes were achieved. All participants completed the three subtasks under the same experimental procedure, allowing task-sensitive cognitive load and multimodal sensing responses to be compared across participants.

Each instructional design subtask was allocated 10 min, resulting in a total task duration of 30 min. The 10 min duration was selected to provide sufficient time for completing each design activity while maintaining a standardized observation window for multimodal sensing comparisons. Accordingly, the three consecutive sensing periods were defined as 0–10 min for instructional objective formulation, 10–20 min for learning activity planning, and 20–30 min for learning assessment design. The participant profile and task overview are summarized in Table 1. The proposed sensor-integrated multimodal cognitive load assessment framework combined physiological sensing, eye-tracking sensing, behavioral process logging, a post-task subjective cognitive load scale, and expert evaluations of instructional design quality. During task execution, physiological sensing signals, eye-tracking data, and behavioral interaction logs were synchronously recorded to capture arousal variation, attentional allocation, hesitation behavior, and revision processes during instructional planning activities. After each task, subjective cognitive load ratings were collected, and the corresponding instructional design products were independently evaluated by experts.

The overall multimodal sensing framework is shown in Fig. 1. Subjective cognitive load was measured after each instructional design subtask using a 12-item questionnaire covering four

Table 1
Participant characteristics and task overview.

Item	Description
Participants	72 preservice elementary teachers
Grouping	36 with practicum experience; 36 without practicum experience
Subtask 1	Instructional objective formulation
Subtask 2	Learning activity planning
Subtask 3	Learning assessment design

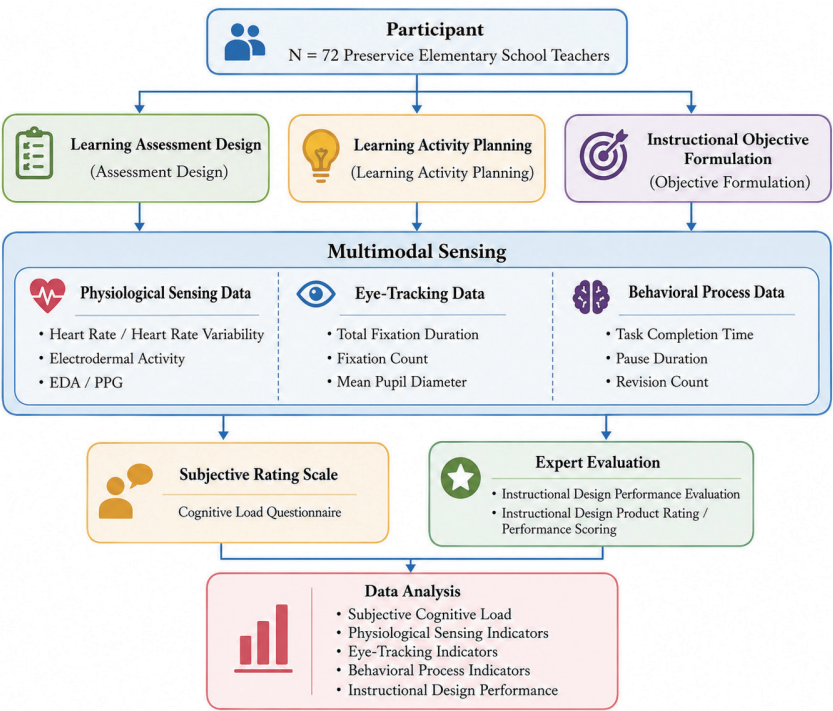


Fig. 1. (Color online) Overall framework of the multimodal sensing assessment for cognitive load in preservice elementary teachers' instructional design tasks.

dimensions: mental effort, perceived difficulty, information integration and decision burden, and overall burden and uncertainty. Each dimension consisted of three items rated on a 7-point Likert scale (1 = strongly disagree to 7 = strongly agree), with higher scores indicating greater perceived cognitive load. The complete questionnaire items and their corresponding dimensions are presented in Table 2. The questionnaire demonstrated good content validity, with a scale-level content validity index (S-CVI) of 0.95. The S-CVI was calculated as the average proportion of questionnaire items judged as content-valid by the expert reviewers, with values closer to 1 indicating stronger overall content validity. The pilot reliability of the three subtask versions ranged from .820 to .899 according to Cronbach's alpha analysis, indicating acceptable-to-good internal consistency. The pilot reliability of the three subtask versions ranged from .820 to .899 according to Cronbach's alpha analysis, indicating acceptable-to-good internal consistency. In addition, instructional design products were evaluated by three experts, and the averaged evaluation results were used as the instructional design quality score.

Table 2

Twelve-item post-task questionnaire for subjective cognitive load assessment.

Item	Dimension	Questionnaire statement
1	Mental effort	I had to invest considerable mental effort to complete this subtask.
2	Mental effort	Completing this subtask required a high level of concentration.
3	Mental effort	I needed to think intensively while completing this subtask.
4	Perceived difficulty	I found this subtask difficult to complete.
5	Perceived difficulty	The requirements of this subtask were complex for me.
6	Perceived difficulty	I experienced difficulty in determining how to complete this subtask appropriately.
7	Information integration and decision burden	I had to integrate multiple pieces of information while completing this subtask.
8	Information integration and decision burden	I had to make several instructional decisions simultaneously during this subtask.
9	Information integration and decision burden	Coordinating instructional objectives, learning activities, and assessment requirements imposed a substantial cognitive burden.
10	Overall burden and uncertainty	I felt uncertain about whether my instructional design decisions were appropriate.
11	Overall burden and uncertainty	I experienced an overall sense of cognitive burden while completing this subtask.
12	Overall burden and uncertainty	I frequently reconsidered or revised my decisions because I was uncertain about the quality of my design.

Note: Each item is rated on a 7-point Likert scale (1 = strongly disagree to 7 = strongly agree). Higher scores indicate greater perceived cognitive load. The 12 items are organized into four dimensions, with three items per dimension.

Instructional design quality was independently evaluated by three experts with experience in elementary teacher education, curriculum and instruction, and instructional design. The experts were purposively selected on the basis of their relevant teaching and research experience. Each instructional design product was evaluated using a 7-point rating scale (1 = very poor to 7 = excellent) covering three criteria: alignment with instructional objectives, appropriateness and coherence of learning activities, and alignment and effectiveness of assessment design. The ratings of the three experts were averaged to obtain the final instructional design quality score for each participant. Nine objective multimodal sensing indicators were used in this study, covering physiological sensing, eye-tracking sensing, and behavioral process logging. The wearable physiological sensing module was used to acquire heart rate, heart rate variability, and electrodermal activity (EDA)-related signals, whereas the eye-tracking module recorded fixation duration, fixation count, and pupil diameter. The behavioral sensing module recorded task completion time, pause duration, and revision frequency through interaction logs. These sensing streams were timestamp-synchronized and subsequently integrated for multimodal cognitive-state analysis. Their operational meanings and cognitive load directions are summarized in Table 3. Physiological signals were acquired using a Shimmer3R GSR+ unit (Shimmer, Ireland), which supports EDA and photoplethysmographic (PPG) measurements for physiological monitoring. Eye-tracking data were collected using a Tobii Pro Fusion eye tracker (Tobii AB, Sweden) to obtain fixation duration, fixation count, and pupil-diameter measurements.

Eye-tracking sensing indicators included fixation duration, fixation count, and pupil diameter for monitoring visual attention allocation and cognitive effort during instructional design

Table 3
Multimodal indicators and operational meanings.

Modality	Indicator	Operational meaning	Load direction
Physiological	Mean heart rate	Average beats per minute during a subtask	Higher = higher load
Physiological	Heart rate variability	Variation in inter-beat intervals	Lower = higher load
Physiological	EDA/pulse-related features	Electrodermal or pulse-response change during a subtask	Higher = higher load
Eye tracking	Total fixation duration	Total valid fixation time on the subtask interface	Higher = higher load
Eye tracking	Fixation count	Number of valid fixations during the task	Higher = higher load
Eye tracking	Mean pupil diameter	Average pupil size during the subtask	Higher = higher load
Behavioral	Task completion time	Time from task start to task submission	Higher = higher load
Behavioral	Mean pause duration	Average duration of no-input or inactive periods	Higher = higher load
Behavioral	Revision count	Number of deletions, rewrites, or adjustments	Higher = higher load

Note: In the “Load direction” column, “=” denotes “indicates” rather than mathematical equality; thus, “Higher = higher load” indicates a positive association with cognitive load, whereas “Lower = higher load” indicates an inverse association.

execution. Behavioral sensing indicators included task completion time, pause duration, and revision frequency to represent interaction efficiency, hesitation behavior, and temporal task progression. Before formal analysis, sensing data were preprocessed through noise reduction, outlier removal, timestamp synchronization, and task-based segment alignment to improve sensing stability and cross-modal consistency. Raw sensing values were retained for statistical analysis, whereas transformed load ratings were additionally generated when necessary for cross-task comparison. The experimental configuration of the multimodal sensing system is shown in Fig. 2. The statistical analysis followed four main stages.

First, repeated-measures analyses were conducted to examine subtask differences in subjective cognitive load across the three instructional design subtasks. Second, descriptive statistical analyses of multimodal sensing indicators were used to compare physiological sensing responses, eye-tracking patterns, and behavioral interaction characteristics across subtasks. Third, independent-samples t-tests were conducted to compare practicum-experience and non-practicum-experience groups in subjective cognitive load and instructional design quality. Fourth, hierarchical clustering analysis was performed as an exploratory multimodal sensing analysis to identify learner cognitive load profiles and heterogeneous cognitive response patterns. Additional analyses, including full correlation matrices and experience-duration comparisons, were treated as secondary analyses supplementary to the primary multimodal sensing investigation. For clarity, the F statistic was used to test overall differences among the three instructional design subtasks in the repeated-measures analysis, whereas the t statistic was used to test differences between the practicum and non-practicum groups. Cohen’s d was calculated as the standardized effect size for group differences. For exploratory clustering, R^2 represented the proportion of variation explained by the cluster solution. The Akaike information criterion (AIC) and Bayesian information criterion (BIC) were used to compare model fit, with lower values indicating a preferable solution, whereas the silhouette coefficient evaluated cluster

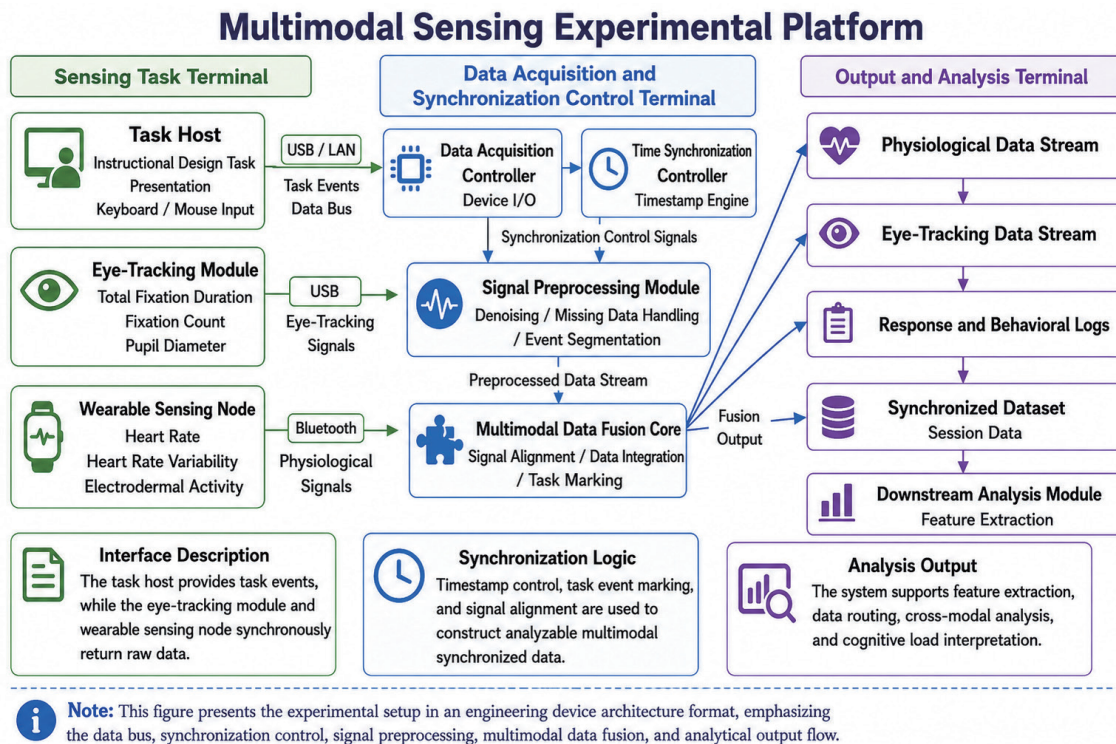


Fig. 2. (Color online) Experimental configuration of the multimodal sensing system.

cohesion and separation, with higher values indicating better clustering quality. Statistical significance was determined at $p < .05$.

3. Results

Significant differences were observed among the three instructional design subtasks across all four subjective cognitive load dimensions. In all dimensions, learning assessment design showed the highest mean cognitive load, followed by instructional objective formulation, whereas learning activity planning showed the lowest load level. These findings indicate that different instructional design subtasks generated distinct cognitive sensing characteristics and task-sensitive cognitive processing demands. For mental effort, the mean scores were 4.986 for instructional objective formulation, 4.727 for learning activity planning, and 5.370 for learning assessment design, with a significant subtask effect, $F = 20.555$, $p < .001$. For perceived difficulty, the corresponding means were 4.991, 4.648, and 5.380, respectively, with $F = 24.977$, $p < .001$. For information integration and decision burden, the means were 4.898, 4.676, and 5.449, respectively, with $F = 32.550$, $p < .001$. For overall burden and uncertainty, the means were 4.921, 4.676, and 5.380, respectively, with $F = 28.653$, $p < .001$. The strongest subtask effect was observed in information integration and decision burden, suggesting that learning assessment design imposed the greatest integrative and decision-related cognitive demand on participants.

From the perspective of multimodal educational sensing, these results indicate that assessment design generated relatively higher cognitive-state variation and mental workload responses during instructional planning activities. From a sensor-development perspective, the framework also identifies practical requirements for educational sensing devices, including long-term wearing comfort, stable skin–sensor contact, low-motion-artifact physiological acquisition, and compact multimodal integration. These requirements may guide the future development of flexible electrodes, skin-compatible sensing materials, and miniaturized wearable sensors for continuous cognitive-state monitoring in educational environments. Overall, the findings demonstrated a stable task pattern in which learning assessment design represented the most cognitively demanding instructional subtask, whereas learning activity planning showed the lowest overall cognitive load level. The corresponding results are summarized in Table 4. Descriptive statistics were further examined to summarize the distributions of major multimodal sensing indicators and instructional design quality across the three instructional design subtasks.

The descriptive statistics of the multimodal sensing indicators and instructional design quality across the three instructional design subtasks are summarized in Table 5. Owing to the large number of sensing parameters and statistical measures, the table is divided into two parts for readability. Part 1 presents the physiological and eye-tracking sensing indicators, whereas Part 2 presents the behavioral sensing indicators and instructional design quality. As summarized in Table 5, the tasks differed in mean heart rate, heart rate variability, EDA-related features, total fixation duration, fixation count, mean pupil diameter, task completion time, mean pause duration, revision count, and instructional design quality. Overall, learning assessment design tended to show relatively higher values across most physiological sensing, eye-tracking sensing, and behavioral interaction sensing indicators, whereas learning activity planning showed the highest mean instructional design quality. These findings suggest that the three instructional design subtasks differed not only in subjective cognitive perception but also in their objective multimodal sensing profiles and instructional performance outcomes.

In particular, higher physiological sensing responses, increased visual attention allocation, and more intensive behavioral interaction patterns were observed during learning assessment

Table 4
Task-sensitive differences in subjective cognitive load across the three instructional design subtasks.

Dimension	Objective formulation, <i>M (SD)</i>	Activity planning, <i>M (SD)</i>	Assessment design, <i>M (SD)</i>	<i>F</i>	<i>p</i>	Order
Mental effort	4.986 (0.717)	4.727 (0.653)	5.370 (0.653)	20.555	< .001	Assessment > Objective > Activity
Perceived difficulty	4.991 (0.712)	4.648 (0.565)	5.380 (0.690)	24.977	< .001	Assessment > Objective > Activity
Information integration and decision burden	4.898 (0.752)	4.676 (0.695)	5.449 (0.683)	32.550	< .001	Assessment > Objective > Activity
Overall burden and uncertainty	4.921 (0.766)	4.676 (0.595)	5.380 (0.654)	28.653	< .001	Assessment > Objective > Activity

Table 5

Descriptive statistics of physiological and eye-tracking sensing indicators across the three instructional design subtasks (Part 1).

Subtask	<i>n</i>	Mean heart rate, <i>M</i> (<i>SD</i>)	Heart rate variability, <i>M</i> (<i>SD</i>)	EDA/pulse-related features, <i>M</i> (<i>SD</i>)	Total fixation duration, <i>M</i> (<i>SD</i>)	Fixation count, <i>M</i> (<i>SD</i>)
Instructional objective formulation	72	4.708 (0.759)	4.458 (0.887)	4.694 (0.725)	4.639 (0.877)	4.639 (0.827)
Learning activity planning	72	4.528 (0.750)	4.139 (0.877)	4.486 (0.712)	4.458 (0.786)	4.278 (0.892)
Learning assessment design	72	4.958 (0.926)	4.708 (0.701)	4.778 (0.755)	4.778 (0.892)	4.833 (0.872)

Descriptive statistics of behavioral sensing indicators and instructional design quality across the three instructional design subtasks (Part 2).

Subtask	Mean pupil diameter, <i>M</i> (<i>SD</i>)	Task completion time, <i>M</i> (<i>SD</i>)	Mean pause duration, <i>M</i> (<i>SD</i>)	Revision count, <i>M</i> (<i>SD</i>)	Mean instructional design quality, <i>M</i> (<i>SD</i>)	Mean pupil diameter, <i>M</i> (<i>SD</i>)
Instructional objective formulation	4.667 (0.650)	4.611 (0.779)	4.611 (0.897)	4.667 (0.822)	5.802 (0.500)	4.667 (0.650)
Learning activity planning	4.611 (0.662)	4.653 (0.790)	4.444 (0.948)	4.306 (0.799)	6.066 (0.434)	4.611 (0.662)
Learning assessment design	4.958 (0.740)	4.806 (0.866)	4.875 (0.838)	4.833 (0.888)	5.688 (0.555)	4.958 (0.740)

design subtasks, indicating elevated cognitive workload and decision-making complexity during assessment construction processes. To provide a more intuitive comparison across sensing modalities, the four subjective cognitive load dimensions were averaged to form a subjective cognitive load index, whereas physiological, eye-tracking, and behavioral sensing indicators were averaged within their respective modalities. As illustrated in Fig. 3, the mean scores are presented as independent comparisons across the four load types rather than as a continuous relationship among them. Learning assessment design showed relatively higher values across all four load types, whereas learning activity planning generally showed lower values. This multimodal sensing pattern was generally consistent with the descriptive results summarized in Table 4 and further supports the effectiveness of the proposed sensor-integrated multimodal cognitive load assessment framework for identifying task-sensitive cognitive processing characteristics during instructional design activities.

To further illustrate temporal sensing variations during instructional design execution, the real-time multimodal sensing responses across the three instructional design subtasks were visualized using synchronized physiological sensing, eye-tracking sensing, and behavioral interaction sensing signals. As illustrated in Fig. 4, learning assessment design consistently produced relatively higher physiological sensing responses, visual attention demand, and behavioral interaction intensity during the later stages of task execution. In particular, heart rate,

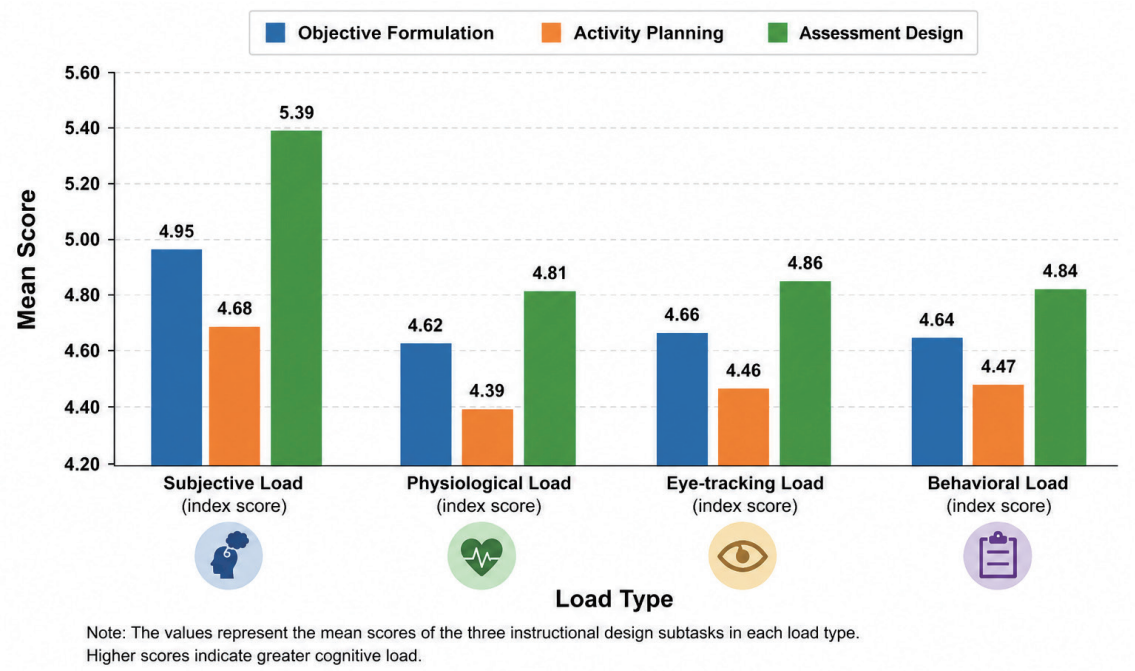


Fig. 3. (Color online) Comparison of subjective cognitive load and multimodal sensing responses across the three instructional design subtasks.

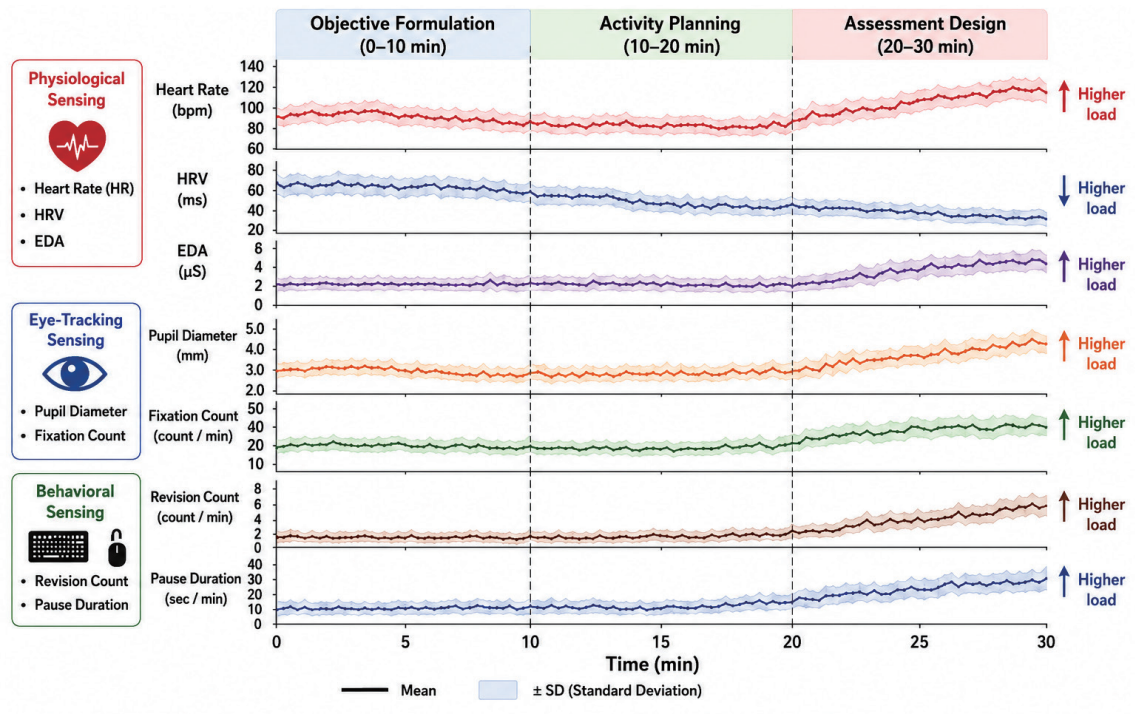


Fig. 4. (Color online) Temporal variations in physiological, eye-tracking, and behavioral sensing responses across three consecutive 10 min instructional design subtasks.

EDA, pupil diameter, fixation count, revision frequency, and pause duration gradually increased during the assessment design stage, whereas heart rate variability showed a decreasing tendency associated with elevated cognitive workload. By comparison, learning activity planning generally demonstrated relatively lower multimodal sensing responses across most sensing dimensions. These temporal sensing patterns further support the effectiveness of the proposed sensor-integrated multimodal cognitive load assessment framework for continuous cognitive-state monitoring and dynamic instructional design analysis.

Participants without practicum experience showed significantly higher subjective cognitive load totals across all three instructional design subtasks, whereas participants with practicum experience consistently demonstrated higher instructional design quality across subtasks. These findings suggest that practicum experience may contribute to more efficient cognitive processing and reduced mental workload during sensor-monitored instructional design activities. For instructional objective formulation, the non-practicum group showed a higher subjective cognitive load total ($M = 64.806$) than the practicum group ($M = 53.972$), $t = 8.410$, $p < .001$, Cohen's $d = 1.982$. For learning activity planning, the corresponding means were 59.222 and 53.139, respectively, $t = 4.748$, $p < .001$, Cohen's $d = 1.119$. For learning assessment design, the non-practicum group again demonstrated a higher subjective cognitive load total ($M = 68.778$) than the practicum group ($M = 60.694$), $t = 6.465$, $p < .001$, Cohen's $d = 1.524$. In contrast, the practicum group consistently achieved higher instructional design quality scores.

For instructional objective formulation, the practicum group scored 5.972, compared with 5.632 in the non-practicum group, $t = -3.052$, $p = .003$, Cohen's $d = -0.719$. For learning activity planning, the means were 6.271 and 5.861, respectively, $t = -4.520$, $p < .001$, Cohen's $d = -1.065$. For learning assessment design, the practicum group scored 5.944, compared with 5.431 in the non-practicum group, $t = -4.412$, $p < .001$, Cohen's $d = -1.040$. Most objective multimodal sensing indicators did not show statistically significant differences between the two groups. From the perspective of multimodal educational sensing, these findings suggest that practicum experience was more strongly associated with reduced subjective cognitive burden and improved instructional performance than with stable differences in physiological sensing responses, eye-tracking behaviors, or behavioral interaction sensing patterns. Therefore, subjective cognitive perception and objective sensing responses may reflect partially different aspects of cognitive processing during instructional planning activities. The group differences in subjective cognitive load and instructional design quality are summarized in Table 6. Hierarchical clustering analysis was further conducted as an exploratory multimodal sensing analysis to identify learner cognitive load profiles and heterogeneous cognitive response characteristics.

Table 6
Practicum group differences in subjective cognitive load and instructional design quality.

Variable	No practicum M	Practicum M	t	p	Cohen's d
T1 subjective total	64.806	53.972	8.410	<.001	1.982
T2 subjective total	59.222	53.139	4.748	<.001	1.119
T3 subjective total	68.778	60.694	6.465	<.001	1.524
T1 design quality	5.632	5.972	-3.052	.003	-0.719
T2 design quality	5.861	6.271	-4.520	<.001	-1.065
T3 design quality	5.431	5.944	-4.412	<.001	-1.040

Among the tested clustering solutions, the 5-cluster model demonstrated the best fit indices, with $R^2 = 0.410$, $BIC = 316.490$, and a silhouette coefficient of 0.190. By comparison, the 3-cluster model showed $R^2 = 0.160$, $BIC = 362.250$, and a silhouette coefficient of 0.140, whereas the 4-cluster model showed $R^2 = 0.181$, $BIC = 376.270$, and a silhouette coefficient of 0.120. Although the 5-cluster solution demonstrated the best overall clustering performance, several clusters were relatively small (e.g., $n = 3$ and $n = 1$). Therefore, the clustering results should be interpreted as exploratory rather than definitive. Nevertheless, the cluster means revealed heterogeneous learner sensing patterns, including low-load/high-performance tendencies, high-load/low-performance tendencies, and other mixed or extreme multimodal cognitive response profiles. These findings support the potential of multimodal sensing features for identifying distinct learner cognitive-state patterns during instructional design subtasks. The clustering results further suggest that multimodal educational sensing technologies may support adaptive cognitive monitoring, learner-state identification, and intelligent instructional support applications in smart teacher education environments. The model comparison and cluster summary are presented in Table 7, whereas the learner cognitive load profiles derived from multimodal sensing indicators are illustrated in Fig. 5.

Figure 5 visualizes the five learner clusters as multidimensional cognitive-load profiles using a radar chart. The five axes represent subjective cognitive load, physiological load, eye-tracking load, behavioral process load, and instructional design quality. The numerical scale represents standardized scores, with 0 indicating the overall sample mean, positive values indicating values above the sample mean, and negative values indicating values below the sample mean for each dimension. Each symbol represents the mean standardized score of a learner cluster on one dimension, and symbols belonging to the same cluster are connected to form its overall multimodal profile. The solid and dotted lines are used only to distinguish the five clusters visually; the connections between adjacent symbols do not indicate temporal, causal, or correlational relationships. Thus, the shape of each profile should be interpreted by comparing its relative position across the five dimensions rather than by interpreting the connecting lines themselves.

4. Discussion

The results consistently demonstrated that learning assessment design imposed the highest subjective cognitive load across all four dimensions. From the perspective of multimodal

Table 7
Exploratory clustering summary of learner cognitive load profiles.

Model	R^2	AIC	BIC	Silhouette	Cluster sizes	Interpretation
3-cluster	0.160	328.100	362.250	0.140	63, 5, 4	Simpler solution but highly unbalanced
4-cluster	0.181	330.740	376.270	0.120	63, 5, 3, 1	Lower silhouette; single-case cluster
5-cluster	0.410	259.570	316.490	0.190	29, 5, 34, 3, 1	Best fit; exploratory due to very small clusters

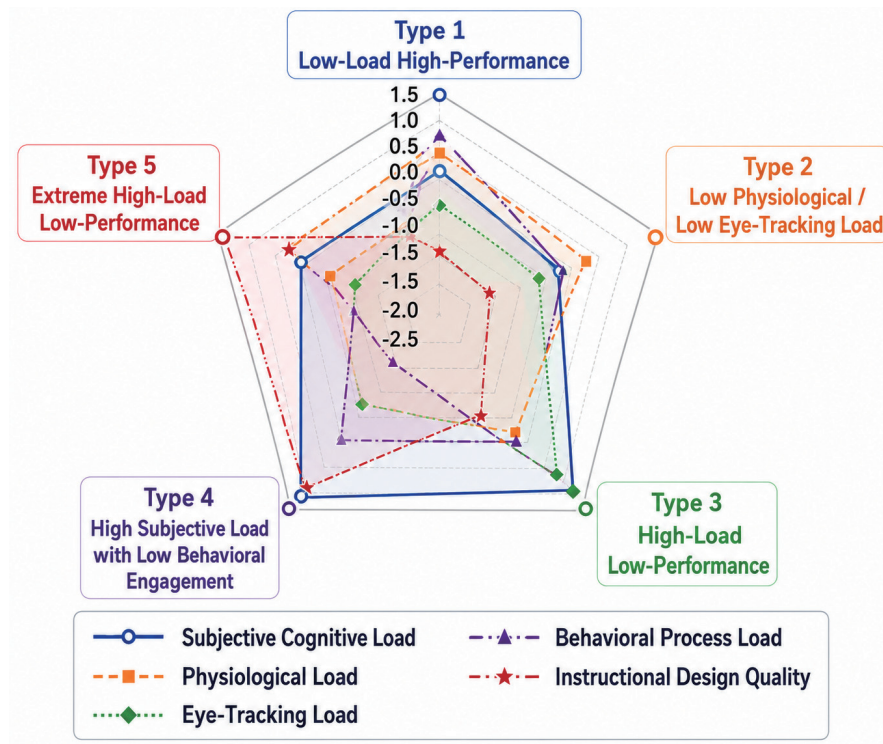


Fig. 5. (Color online) Multimodal learner cognitive-load profiles identified through exploratory clustering analysis. Values are standardized scores, where 0 represents the overall sample mean, positive values indicate above-average values, and negative values indicate below-average values for each dimension. Symbols connected by the same line represent the multidimensional profile of one learner cluster; the connecting lines are used only for visualization and do not indicate temporal, causal, or correlational relationships.

educational sensing, this finding suggests that assessment construction generated relatively stronger cognitive-state fluctuations, physiological arousal responses, visual attention demands, and behavioral interaction during instructional planning activities. A plausible explanation is that assessment design requires continuous backward referencing to instructional objectives, alignment verification with learning activities, and the generation of valid evidence for evaluating student learning outcomes. This process involves repeated information integration, evidence-based reasoning, and instructional decision-making, thereby increasing cognitive workload and mental processing complexity. These characteristics may also explain why the information integration and decision burden dimension demonstrated the strongest subtask effect among the four cognitive load dimensions. The proposed sensor-integrated multimodal cognitive load assessment framework successfully captured two important forms of variation during instructional design execution. First, the framework differentiated task-sensitive cognitive demands because the three instructional design subtasks demonstrated distinct subjective cognitive load patterns and heterogeneous multimodal sensing responses.

Learning assessment design consistently generated relatively higher physiological sensing responses, eye-tracking sensing indicators, and behavioral interaction sensing patterns compared with the other instructional design subtasks. Second, the framework successfully differentiated learner-experience characteristics because participants with practicum experience consistently reported lower subjective cognitive load and achieved higher instructional design quality. These

findings indicate that the proposed multimodal sensing framework was sensitive to both task-level cognitive variation and learner-level performance differences during instructional planning activities. The present study also found limited correspondence between subjective cognitive load totals and objective multimodal sensing indicators. However, this result should not be interpreted as measurement inconsistency or sensing failure. Instead, it suggests that subjective cognitive perception and objective sensing responses may represent different layers of cognitive processing during instructional design execution. Subjective cognitive load ratings summarize learners' overall perceptions of difficulty, mental effort, uncertainty, and task burden after task completion, whereas objective sensing indicators continuously capture physiological arousal, visual attention allocation, and behavioral hesitation during task execution.

Therefore, subjective evaluations and multimodal sensing responses should be interpreted as complementary rather than contradictory sources of cognitive-state information. From the perspective of multimodal educational sensing, the integration of heterogeneous sensing signals may provide more comprehensive cognitive monitoring capability than reliance on single-source measurements alone. The findings of this study have important implications for smart teacher education and sensor-assisted instructional support systems. The proposed multimodal sensing framework demonstrates the feasibility of integrating physiological sensing, eye-tracking sensing, behavioral interaction sensing, and synchronized sensor fusion analytics into instructional design training environments. Such sensing systems may support real-time cognitive overload warning, adaptive instructional prompting, learner-state identification, task-stage diagnosis, and intelligent instructional assistance in future smart teacher education applications. Furthermore, the proposed framework may provide a foundation for adaptive educational sensing systems capable of continuously monitoring learner cognitive states and dynamically adjusting instructional support strategies during instructional planning activities.

Several limitations should nevertheless be acknowledged. First, the sample was collected from a single institution, which may limit the generalizability of the findings across different educational environments and teacher-training contexts. Second, the associations between subjective cognitive load and objective sensing indicators were relatively limited, suggesting that additional multimodal sensing features and improved cross-modal integration methods may be necessary to enhance sensing sensitivity and cognitive-state interpretation capability. Third, the exploratory clustering analysis included several relatively small learner clusters, reducing clustering stability and limiting the interpretability of some multimodal learner profiles. Future studies should therefore include larger and more diverse participant samples, improve multimodal sensor-fusion algorithms, incorporate additional sensing modalities, and further validate the generalizability of multimodal learner cognitive profiles across different instructional design contexts and smart educational sensing environments.

5. Conclusions

In this study, we developed a sensor-integrated multimodal cognitive load assessment framework for preservice elementary teachers' instructional design tasks by integrating physiological sensing, eye-tracking, behavioral process logging, subjective cognitive load ratings, and expert evaluations. The results revealed significant differences among the three instructional design subtasks, with learning assessment design producing the highest cognitive

demand and learning activity planning the lowest. The framework also identified practicum-related differences in cognitive load and instructional performance and revealed heterogeneous learner cognitive profiles. Temporal analysis further demonstrated dynamic variations in physiological, eye-tracking, and behavioral sensing responses across instructional design stages. These findings support the potential of multimodal sensing for cognitive overload warning, learner-state monitoring, adaptive scaffolding, and intelligent teacher education. Beyond educational analytics, the framework provides application requirements for future wearable sensors, particularly flexible and skin-compatible sensing materials and compact multimodal integration. The main limitations include the single-institution sample, limited correspondence between subjective and objective sensing indicators, and exploratory learner profiling. Remaining challenges involve improving cross-modal sensor fusion and establishing robust relationships between multimodal sensing responses and cognitive states. Future studies should validate the framework with larger and more diverse samples and improve multimodal sensing integration for adaptive educational applications.

Acknowledgments

This work was supported by Summit-Tech Resource Corp. and by projects under Nos. NSTC 113-2221-E-390-011 and NSTC 114-2622-E-390-001.

References

- 1 R. W. Tyler: *Basic Principles of Curriculum and Instruction* (University of Chicago Press, Chicago, 2013) original work published in 1949.
- 2 L. W. Anderson and D. R. Krathwohl, Eds.: *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives* (Longman, New York, 2001).
- 3 G. Wiggins and J. McTighe: *Understanding by Design* (Association for Supervision and Curriculum Development, Alexandria, 2005) expanded 2nd ed.
- 4 J. Biggs and C. Tang: *Teaching for Quality Learning at University* (McGraw-Hill Education, Berkshire, 2011) 4th ed.
- 5 F. G. W. C. Paas: *J. Educ. Psychol.* **84** (1992) 429.
- 6 S. G. Hart and L. E. Staveland: *Human Mental Workload*, P. A. Hancock and N. Meshkati, Eds. (North-Holland, Amsterdam, 1988) p. 139.
- 7 J. Leppink, F. Paas, C. P. M. Van der Vleuten, T. Van Gog, and J. J. G. Van Merriënboer: *Behav. Res. Methods* **45** (2013) 1058.
- 8 A. Jimenez-Molina, C. Retamal, and H. Lira: *Sensors* **18** (2018) 458.
- 9 A. M. Hughes, P. A. Hancock, L. R. Marusich, J. L. Szalma, and C. Hoare: *Hum. Factors* **61** (2019) 393.
- 10 B. Thielmann, R. Pohl, and I. Böckelmann: *J. Occup. Med. Toxicol.* **16** (2021) 24.
- 11 V. Skaramagkas, G. Giannakakis, A. Ktistis, D. Manousos, I. Karatzanis, N. Tachos, E. E. Tripoliti, K. Marias, D. I. Fotiadis, and M. Tsiknakis: *IEEE Rev. Biomed. Eng.* **16** (2023) 260.
- 12 H. K. Wong and J. Epps: *Comput. Methods Programs Biomed.* **137** (2016) 47.
- 13 A. M. Harris, J. O. Eayrs, and N. Lavie: *Cogn. Res. Princ. Implic.* **8** (2023) 56.
- 14 F. C. Specian Junior, D. Litchfield, J. Sandars, and D. Cecilio-Fernandes: *Med. Teach.* **46** (2024) 1502.
- 15 A. Heilmann and S. Neumann: *Proc. Workshop on Computational Linguistics for Linguistic Complexity (CL4LC)* (The COLING 2016 Organizing Committee, Osaka, 2016) p. 98.
- 16 F. N. Biondi, A. Cacanindin, C. Douglas, and J. Cort: *Hum. Factors* **63** (2021) 813.
- 17 Y. Wang: *Appl. Sci.* **15** (2025) 4401.