Employing Hierarchical Cluster Analysis to Investigate the Modality of Digital Teaching in Regular Classrooms

Pei-Lun Wang [1], Yung-Hsiang Huang [2], Min-Hsien Lee [3], Chia-Ching Lin [4*]

[1] National Kaohsiung Normal University

[2] National Kaohsiung Normal University

[3] National Taiwan Normal University

[4] National Kaohsiung Normal University

[*] mcclin@mail.nknu.edu.tw

Abstract: *Recent research has underscored the growing role of digital technologies in classroom teaching, a trend further accelerated by the COVID-19 pandemic and advances in artificial intelligence. This study examines how digital teaching is enacted in everyday classroom practice. Classroom activities of 39 teachers were video-recorded and analyzed to capture pedagogical strategies, digital applications, and representational modes. A total of 22 indicators were identified and analyzed using Hierarchical Cluster Analysis (HCA) to reveal distinct configurations of digital teaching. The findings suggest that effective digital teaching depends not only on the use of technology but also on the coherent integration of instructional strategies, interaction patterns, and representational modes.*

Keywords: Digital Teaching, Multimodality, Instructional Design, Video Analysis, Hierarchical Cluster Analysis

1. Introduction

Over the past decade, the integration of digital technologies into school classrooms has become an increasingly common feature of everyday teaching practice. In Taiwan, large-scale initiatives promoting technology-enhanced instruction have led to widespread access to digital tools such as interactive whiteboards, mobile devices, and online learning platforms. As a result, digital technology is no longer regarded as an innovative add-on but rather as a routine component of classroom environments. In many schools, teachers who participate in professional development programs on digital pedagogy are encouraged to integrate these technologies into their everyday teaching.

While a substantial body of research has examined the effectiveness of technology-enhanced learning environments, much of this work has focused on learning outcomes, student achievement, or specific instructional interventions. Comparatively less attention has been paid to what happens in classrooms where digital technology has already been integrated into routine teaching practice. In other words, once digital tools become part of everyday instruction, an important question emerges what kinds of multimodal teaching configurations take shape in these classrooms (Ertmer & Ottenbreit-Leftwich, 2010). In line with social semiotic perspectives, this study adopts the concept of modes as socially shaped resources for meaning making and focuses on how multiple modes are configured in classroom teaching.

In this study, observable classroom activities are represented through a set of indicators corresponding to different modes, and recurring configurations of these indicators are interpreted as modalities of digital teaching. To address this gap, the present study employs hierarchical cluster analysis (HCA) to uncover underlying patterns of multimodal teaching across 39 technology-enhanced classrooms. Rather than evaluating the efficacy of specific digital tools, we focus on the real-world orchestration of hardware, digital resources, representational modes, interaction structures, and pedagogical arrangements. By identifying distinct multimodal teaching profiles, this study provides a systematic account of how digital pedagogy is actually enacted in everyday instruction. Accordingly, we address the following research questions:

1. What are the key factors related to the deployment of digital teaching in daily classrooms?

2. How is the modality of digital teaching manifested in daily classrooms?

2. Methodology

2.1. Analytic Sample and Classroom Context

The analytic sample consisted of classroom video recordings from 39 in-service teachers in southern Taiwan where digital technologies had been integrated into regular instructional practices. These cases were selected from a corpus of public lessons conducted by teachers participating in Taiwan's nationwide digital learning initiative.

The participating teachers included 10 males and 29 females. Two teachers, one male and one female, were from junior high schools, while the remaining teachers taught in elementary schools. All teachers had completed at least 10 hours of professional development related to digital learning platforms and instructional design. As part of the program requirements, each teacher was expected to integrate digital technologies into their classroom teaching at least once per week throughout the semester and to deliver a public lesson at the end of the semester.

In addition to conventional classroom materials, the participating classrooms were equipped with wireless internet access, large interactive displays, and a one-to-one mobile device environment for both teachers and students. The specific digital learning platforms and tools used during instruction were selected and integrated by the teachers in line with their curricular objectives.

Consequently, the selected cases represent classroom contexts in which both teachers and students had prior experience with technology-integrated instruction. This ensured that the recorded lessons reflected routine classroom practices rather than experimental or demonstration-based teaching situations.

2.2. Analytical Framework

To analyze multimodal teaching practices in technology-enhanced classrooms, an analytical framework was developed to capture different aspects of classroom activity. The analytical framework was informed by multimodal theory, which emphasizes that meaning in classroom interaction is constructed through the coordinated use of multiple semiotic resources such as language, images, gestures, and material tools (Jewitt, 2009; Jewitt et al., 2016; Kress, 2010; Kress & van Leeuwen, 2001; Norris, 2004). Research in science education has also highlighted the importance of examining how different representational resources are combined during teaching and learning processes (Tang, Delgado, & Moje, 2014). Furthermore, the development of the coding categories in this study was informed by recent methodological advances in multimodal video analysis, which emphasize the systematic integration of visual, verbal, and interactive cues to examine technology-enhanced learning behaviors (Serafini & Reid, 2023; Yusuf et al., 2024).

Classroom videos were segmented into one-minute intervals, which served as the unit of analysis. Within each interval, the presence of specific multimodal indicators was identified and recorded using a binary system in which a value of 1 indicated that a particular indicator was observed and a value of 0 indicated that it was not present.

The coding framework consists of five dimensions: pedagogical modes, hardware and physical carriers, digital resources, representational modes, and interaction agents. These dimensions capture key aspects of classroom activity, including how instruction is organized, what technological and material resources are used, how information is represented, and how interaction occurs among teachers, students, and digital systems. The detailed coding categories used in the analysis are summarized in Table 1.

Table 1.

Coding framework of digital teaching in practice

Dimension	Codes	Description

Pedagogical modes	Guided learning	Teacher-directed instruction where the teacher explains concepts or demonstrates procedures
	Collaborative learning	Students work together to complete tasks or solve problems
	Peer learning	Students explain ideas or provide feedback to peers
	Self-regulated learning	Students work independently with minimal teacher guidance
Hardware and physical carriers	Interactive display	Interactive whiteboards or projection systems used during instruction
	Mobile devices	Tablets or smartphones used for learning activities
	Computers	Desktop or laptop computers used in classroom instruction
	Physical apparatus	Experimental tools or physical teaching aids
Digital resources	Multimedia content	Videos, animations, or other multimedia materials used in teaching
	Digital documents	Digital texts, slides, worksheets, or presentation materials
	Simulations	Interactive simulations illustrating scientific phenomena
	Learning platforms	Online platforms used for instruction or interaction
Representational modes	Spoken language	Verbal explanations by teachers or students
	Written text	Textual information displayed on screen or board
	Static visual representations	Diagrams, graphs, or images that remain visually stable
	Dynamic visual representations	Moving visual representations such as animations or videos
	Gestures and embodied actions	Body movements used to illustrate ideas or guide attention
Interaction agents	Teacher-led interaction	Communication primarily directed by the teacher
	Teacher–student interaction	Direct exchange between teacher and students
	Student–student interaction	Communication among students during learning activities
	Student–technology interaction	Interaction between students and digital systems

2.3. Video Coding Procedure

The coded categories were subsequently transformed into binary codes for cluster analysis. This study employed video analysis as the primary analytical method. Each video was segmented into one-minute intervals, with each interval serving as the unit of analysis. For example, a 40-minute lesson would yield 40 coding units. Each unit was analyzed using the analytical framework described earlier. A binary coding approach was adopted in which the presence of a particular multimodal element within a coding unit was assigned a value of 1, whereas its absence was assigned a value of 0.

2.4. Inter-Rater Reliability

Three independent coders were recruited for the video coding process. The 39 classroom videos were randomly assigned to the coders for independent analysis. Prior to formal coding, all coders participated in a training session to ensure a shared understanding of the definitions and criteria associated with each indicator. To establish coding reliability, the three coders independently coded a shared subset of practice data, and inter-rater agreement was assessed using Krippendorff's alpha (α = .64). Training and discussion before formal coding represent a commonly adopted approach for establishing reliability in content analysis (Lombard et al., 2002). After the training phase, the remaining video data were distributed among the coders and coded independently.

Given the high complexity of the 22-dimensional multimodal coding scheme and the dynamic, concurrent nature of real-world classroom interactions, the obtained alpha value ($\alpha = .64$) was considered acceptable for this exploratory research (Krippendorff, 2019).

2.5. Data Analysis

To identify patterns of multimodal teaching practices across classrooms, hierarchical cluster analysis (HCA) was conducted on the coded instructional indicators. The analysis was performed using Ward's clustering method (Ward.D2) with Euclidean distance as the similarity measure. Each classroom case was represented by a set of operationalized indicators derived from the analytical framework, allowing similarities in instructional configurations to be examined across the 39 lessons. The optimal number of clusters was determined through inspection of the dendrogram together with considerations of interpretability and cluster stability. In this study, the operationalized indicators represent different modes of classroom activity, and the clusters derived from HCA are interpreted as distinct modalities of digital teaching.

After the clusters were identified, a permutational multivariate analysis of variance (PERMANOVA; Anderson, 2001) was conducted to examine whether the clusters differed significantly in their overall multimodal profiles. The analysis was performed using a Bray–Curtis dissimilarity matrix with 9,999 permutations. Prior to the PERMANOVA, PERMDISP was applied to verify the homogeneity of multivariate dispersion.

To further examine which specific instructional indicators contributed to differences among clusters, nonparametric Kruskal–Wallis tests were conducted for each coded variable. When significant differences were detected, post hoc pairwise comparisons were performed using the Dwass–Steel–Critchlow–Fligner (DSCF) procedure to identify differences between clusters.

3. Results

3.1. Identification of Multimodal Teaching Profiles

Hierarchical cluster analysis was conducted to identify patterns of multimodal teaching practices across the 39 classroom cases. Inspection of the dendrogram suggested a three-cluster solution that provided an interpretable grouping of instructional configurations.

Cluster 1 represents a collaborative multimodal configuration in which students frequently engage in peer interaction and group-based activities supported by a range of representational resources.

Cluster 2 reflects a teacher-guided multimodal configuration characterized by a predominance of teacher-led instruction, supported by interactive display technologies and spoken explanations.

Cluster 3 represents a technology-mediated multimodal configuration, in which students frequently interact with digital devices and learning platforms, with a stronger emphasis on student–technology interaction and digitally mediated representations.

3.2. Overall Multivariate Differences Across Clusters

The PERMDISP result showed no significant differences in within-group dispersion across clusters ($F[2,36] = 3.54$, $P>.05$), thus satisfying the assumption underlying PERMANOVA. The overall PERMANOVA yielded a significant three-cluster solution ($F[2,36]=8.61$, $p<.001$, $R^2=.324$), with cluster membership accounting for 32.4% of the total multivariate variance across the 22 dimensions with a large effect, indicating significant differences among 3 clusters. Pairwise post-hoc comparisons with Bonferroni correction ($\alpha=.017$) further revealed significant differences between all cluster pairs: the Collaborative vs. the Teacher-Guided ($F[1,34]=8.04$, $p<.001$, $R^2=.191$), the Collaborative vs. the Technology-Mediated ($F[1,11]=6.16$, $p<.05$, $R^2=.359$), and the Teacher-Guided vs. Technology-Mediated ($F[1,27]=11.31$, $p<.01$, $R^2=.295$).

Following the establishment of overall multivariate group differences, Kruskal-Wallis nonparametric tests were conducted to identify specific differences across individual indicators. Of the 22 indicators, 11 reached statistical significance ($p<.05$), all with effect sizes at or above the moderate threshold ($\varepsilon^2 \geq .14$). Significant indicators spanned all five pedagogical domains. Organized by domain, these were: in the instructional context domain (P), P1 ($\varepsilon^2=.561$) and P3 ($\varepsilon^2=.273$); in the hardware domain (H), H1 ($\varepsilon^2=.532$) and H2 ($\varepsilon^2=.187$); in the digital resources domain (D), D5 ($\varepsilon^2=.313$) and D3 ($\varepsilon^2=.240$); in the representational modalities domain (R), R1 ($\varepsilon^2=.578$) and R5 ($\varepsilon^2=.223$); and in the interaction domain (I), I3 ($\varepsilon^2=.998$), I1 ($\varepsilon^2=.471$), and I4 ($\varepsilon^2=.385$). Further, the DSCF results were illustrated alongside the indicators on the radar chart (Figure 1). It should be noted that statistical comparisons involving the third cluster were interpreted cautiously and supplemented with descriptive analysis since the third cluster contains a relatively small number of cases (n=3).

The descriptive statistics of the multimodal indicators across the three clusters are summarized in Table 2.

Table 2.

Descriptive statistics of DTP indicators across clusters

Indicator	Cluster 1 Mean (S.D.)	Cluster 2 Mean (S.D.)	Cluster 3 Mean (S.D.)
P1 Guided learning	0.29(0.11)	0.58(0.13)	0.5(0.08)
P2 Collaborative learning	0.29(0.21)	0.16(0.13)	0.26(0.18)
P3 Peer learning	0.33(0.22)	0.13(0.13)	0.08(0.07)
P4 Self-regulated learning	0.08(0.10)	0.13(0.17)	0.16(0.15)
H1 Interactive display	0.34(0.16)	0.66(0.18)	0.05(0.06)
H2 Mobile devices	0.14(0.19)	0.18(0.31)	0.78(0.17)
H3 Computers	0.03(0.07)	0.07(0.15)	0(0.00)
H5 Physical apparatus	0.06(0.06)	0.14(0.17)	0.43(0.43)
D1 Multimedia content	0(0.00)	0.01(0.05)	0(0.00)
D3 Digital documents	0.12(0.13)	0.33(0.27)	0.71(0.16)
D4 Simulation	0.03(0.19)	0.08(0.16)	0.02(0.03)
D5 Learning platform	0.19(0.11)	0.43(0.25)	0.03(0.06)
D6 Other digital resources	0.25(0.32)	0.14(0.12)	0.06(0.11)
R1 Spoken language	0.61(0.20)	0.93(0.17)	0.4(0.21)
R2 Written text	0.31(0.41)	0.38(0.42)	0.91(0.08)
R3 Static visual	0.15(0.25)	0.23(0.34)	0.32(0.26)
R4 Dynamic visual	0.03(0.11)	0.03(0.09)	0.02(0.04)
R5 Gesture	0.56(0.29)	0.7(0.20)	0.16(0.12)
I1 Teacher-led interaction	0.41(0.23)	0.76(0.22)	0(0.00)
I2 Teacher–student interaction	0.45(0.31)	0.51(0.27)	0.78(0.01)
I3 Student–student interaction	0(0.00)	0(0.00)	0.38(0.17)
I4 Student–technology interaction	0.32(0.13)	0.64(0.24)	0.8(0.17)

3.3. Multimodal Configuration Profiles

To visualize the multimodal characteristics of the clusters, the mean values of the coded indicators were plotted using a radar chart (Figure 1). The radar chart illustrates how different instructional resources and interaction patterns are configured across the three clusters.

The Collaborative multimodal configuration cluster shows relatively higher values in indicators related to collaborative learning and student interaction. The Teacher-guided multimodal configuration cluster emphasizes guided instruction and teacher-led interaction supported by interactive display technologies. In contrast, the Technology-mediated multimodal configuration cluster shows higher values in indicators related to student–technology interaction and the use of digital learning platforms. This variety suggests that providing identical hardware does not automatically lead to uniform pedagogical shifts. The distinct profiles, particularly the contrast between teacher-guided and collaborative configurations, underscore that teachers' pedagogical orchestration skills determine whether technology acts as a mere substitute for traditional tools or as a catalyst for new interactive learning models.

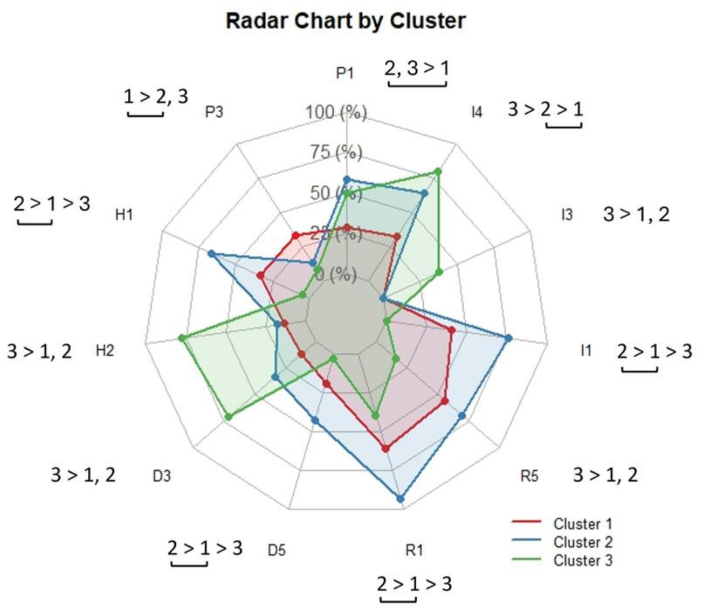

Figure 1

Radar chart of multimodal teaching profiles

4. Discussion and Implications

The results of this study suggest that digital teaching in everyday classrooms does not follow a single instructional model. Instead, technology integration appears to support multiple multimodal teaching configurations that differ in their pedagogical organization, interaction structures, and representational resources. The three clusters identified in this study illustrate how teachers combine technological tools with instructional strategies in different ways. Some classrooms emphasize collaborative learning supported by digital platforms, whereas others rely more heavily on teacher-guided instruction using interactive display technologies. A smaller group of classrooms demonstrates intensive use of mobile devices and learning platforms that facilitate direct student–technology interaction.

These findings highlight the importance of examining how digital technologies are pedagogically orchestrated in classroom practice rather than focusing solely on the availability of technological tools. Understanding these multimodal configurations may help inform teacher professional development programs that aim to support more effective integration of digital technologies in classroom teaching. These findings suggest that teacher professional development should focus not only on technological skills but also on how teachers orchestrate multimodal resources in classroom instruction.

References

Anderson, M. J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecology*, 26(1), 32–46. https://doi.org/10.1111/j.1442-9993.2001.01070.pp.x

Ertmer, P. A., & Ottenbreit-Leftwich, A. T. (2010). Teacher technology change: How knowledge, confidence, beliefs, and culture intersect. *Journal of Research on Technology in Education*, 42(3), 255–284. https://doi.org/10.1080/15391523.2010.10782551

Jewitt, C. (Ed.). (2009). *The Routledge handbook of multimodal analysis*. Routledge.

Jewitt, C., Bezemer, J., & O’Halloran, K. (2016). *Introducing multimodality.* Routledge. https://doi.org/10.4324/9781315638027

Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. Routledge. https://doi.org/10.4324/9780203970034

Kress, G., & van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. Arnold.

Krippendorff, K. (2019). *Content analysis: An introduction to its methodology (4th ed.).* Sage.

Lombard, M., Snyder-Duch, J., & Bracken, C. C. (2002). Content analysis in mass communication: Assessment and reporting of intercoder reliability. *Human Communication Research*, 28(4), 587–604. https://doi.org/10.1111/j.1468-2958.2002.tb00826.x

Norris, S. (2004). *Analyzing multimodal interaction: A methodological framework*. Routledge. https://doi.org/10.4324/9780203379493

Serafini, F., & Reid, S. F. (2023). Multimodal content analysis: Expanding analytical approaches to content analysis. *Visual Communication*, *22*(4), 623–649. https://doi.org/10.1177/1470357219864133

Tang, K. S., Delgado, C., & Moje, E. B. (2014). An integrative framework for the analysis of multiple and multimodal representations for meaning making in science education. *Science Education*, 98(2), 305–326. https://doi.org/10.1002/sce.21099

Yusuf, A., Md Noor, N., & Bello, S. (2024). Using multimodal learning analytics to model students' learning behavior in animated programming classroom. *Education and Information Technologies*, *29*(6), 6947–6990. https://doi.org/10.1007/s10639-023-12079-8