Advances in Medical Education and Practice

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Contextual Vertical Integration Optimizes the Cognitive Load Profile of Undergraduate Clinical Obstetrics and Gynecology Learners: Evidence from the CONVINCED-OG Cluster-Randomized Crossover Trial

Mohammed Ismail-Khan [1,2], Dominic A Blake [1], Mona MI Abdalla [3], Porfyrios Korompelis [1]

[1]Northern Gynaecological Oncology Centre, Queen Elizabeth Hospital, Gateshead, UK; [2]Department of Physiology & Medical Education, Shadan Institute of Medical Sciences, Hyderabad, India; [3]Department of Human Biology, School of Medicine, International Medical University, Kuala Lumpur, Malaysia

Correspondence: Mohammed Ismail-Khan, Northern Gynaecological Oncology Centre, Queen Elizabeth Hospital, Gateshead, NE9 6SX, UK, Email Ismail.khan11@nhs.net

Background: The disjunction between basic sciences and clinical teaching impairs knowledge retention and application. Contextual vertical integration (CVI) embeds basic-science explanations within clinical instruction and may optimize cognitive load by lowering intrinsic cognitive load (ICL) and extraneous cognitive load (ECL), which impede learning and enhance germane cognitive load (GCL) that facilitates learning. The Contextual Vertical Integration in Clinical Education in Obstetrics and Gynecology (CONVINCED-OG) trial evaluates the role of CVI in optimizing the cognitive load.

Methods: This cluster-randomized crossover study involved three year-3 undergraduate medical student cohorts from Newcastle University at two teaching units. Each cohort was taught eight topics (four obstetrics, four gynecology). A computer-generated Latin-square schedule assigned two cohorts to CVI and one to standard teaching for each topic, then rotated the pattern, so every cohort received four CVI and four standard sessions. The same teacher delivered all sessions with identical materials at both sites, with students remaining unaware of their randomization status. Perceived cognitive loads (ICL, ECL and GCL) were self-reported using the validated Klepsch's Cognitive Load Questionnaire. Topic-level differences (integrated vs standard) were analyzed with independent *t*-tests and pooled within-student differences were assessed with paired *t*-tests.

Results: Across all eight topics, CVI reduced ICL and ECL, alongside significantly increasing GCL (except one topic where ICL differences were insignificant). Pooled within-subject analyses corroborated these findings, demonstrating consistent cognitive load optimisation with CVI.

Conclusion: By demonstrating the role of CVI in optimizing cognitive load across diverse topics within undergraduate clinical obstetrics and gynecology education, CONVINCED-OG trial supports the adoption of CVI in clinical teaching.

Keywords: basic sciences, medical education, schema, intrinsic cognitive load, extraneous cognitive load, germane cognitive load

Introduction

Vertical Integration and the Role of Contextuality

Undergraduate medical education often separates the study of basic sciences from clinical instruction. Typically, students acquire foundational knowledge in the basic sciences during the first two to three years of medical school, followed by clinical training in the subsequent two to three years. Graduate outcomes outlined by the Liaison Committee on Medical Education in the United States and Canada,[1] the Australian Medical Council[2] and the General Medical Council in the

https://doi.org/10.2147/AMEP.S557502
Received: 3 August 2025
Accepted: 14 December 2025
Published: 12 January 2026

United Kingdom[3] mandate that graduates apply basic scientific principles in clinical practice, given its proven efficacy in enhancing patient care.[4,5]

Despite this, students often struggle to relate the clinical knowledge acquired in later years to the basic sciences taught earlier. This challenge is largely attributed to the temporal and spatial segregation in basic science and clinical specialties,[6] with studies reporting significant knowledge attrition in the basic sciences of up to 50% within a single - year.[7–9] This disjunction may impair clinical reasoning, limit diagnostic accuracy, and reduce the depth of conceptual understanding required for patient care.[10,11]

Vertical integration, as a pedagogical strategy, seeks to illustrate basic science principles within the clinical context.[12] This endeavors to bridge the gap between basic sciences and clinical knowledge, thereby enhancing the students' ability to engage meaningfully in clinical problem solving.[6] Harden describes the transdisciplinary level of integration as the highest rung on the integration ladder, where disciplines extend beyond the boundaries of academic confines to improve the holistic understanding of a clinical topic.[12,13] For example, understanding polycystic ovarian syndrome requires the integration of reproductive physiology, the biochemistry of androgens, insulin resistance pathophysiology, ovarian sonographic anatomy, and pharmacology related to metformin, estrogen, progestins, and antiandrogens.

Instructional designs that aspire to this level of integration must achieve micro-level vertical integration within each element of an individual clinical topic. This involves linking the "cause" (basic sciences) with the "effect" (clinical manifestation or application) through an explanatory narrative, rather than merely juxtaposing these content areas.[6,13] This specific approach is referred to in this study as contextual vertical integration (CVI), wherein basic science principles are interwoven into the clinical content being taught.

Vertical integration has traditionally been conceptualized as the bidirectional linking of basic and clinical sciences across all stages of medical education, allowing clinical relevance to be introduced early and basic science principles to be revisited during later clinical training.[6,12] The present approach of CVI differs in its micro-instructional embedding of basic science explanations within distinct elements of the clinical topic. Similarly, spiral integration emphasizes revisiting core topics at increasing levels of complexity across the curriculum[6,13] whereas CVI emphasizes real-time, context-triggered science-in-practice within each discrete element of a clinical topic.

However, despite the consensus that integrating basic and clinical sciences enhances learning, most empirical studies have evaluated this relationship through performance-based outcomes such as examination scores, diagnostic reasoning tasks, or knowledge retention.[10,14–17] The underlying learning process, specifically how instructional design influences learners' cognitive architecture during learning, has been largely unexplored. This study addresses that gap by investigating the cognitive effects of CVI-enabled instructional design through the lens of cognitive load theory (CLT).

Cognitive Load Theory and Instructional Design

Sweller's CLT,[18,19] which is grounded in Piaget's schema theory,[20] conceptualizes learning as the process of schema acquisition. CLT provides a robust framework for evaluating how specific instructional strategies influence the allocation of cognitive resources between learning-relevant and learning-irrelevant processes.

CLT is based on the premise that working memory, which plays a pivotal role in schema acquisition, has limited capacity.[21,22] It processes discrete information received from sensory memory into an organized knowledge structure–known as a cognitive schema-and subsequently facilitates its storage in long-term memory, which has limitless capacity.[22,23]

To explain schema acquisition, CLT identifies three categories of cognitive loads:[24] intrinsic cognitive load (ICL), extraneous cognitive load (ECL) and germane cognitive load (GCL). ICL reflects the demands on working memory due to the inherent complexity of the material being taught, which is influenced by the interactivity of the elements of information within that topic, and the learner's prior knowledge.[21,25,26] ECL refers to the demand placed on working memory from poorly designed instruction, including disorganized content, lack of coherence, or ineffective presentation.[21,24] Both the ICL and the ECL can place a "load" which can impede schema acquisition, if it exceeds the limited working memory capacity.[26]

GCL was originally defined as the cognitive effort devoted to facilitating schema acquisition and represented the "load" placed on working memory for learning-relevant activities.[23,24,26] However, in contemporary interpretations of

CLT, GCL is reconceptualized from a distinct category of "load" to the proportion of working memory resources that are directed away from extraneous learning-irrelevant demands towards intrinsic learning-relevant processing, thus enabling schema acquisition.[18,21]

Optimizing instructional design thus involves managing ICL, minimizing ECL and promoting the allocation of working memory resources towards schema-relevant activities (i.e. GCL).[25,26] This study is consistent with Sweller's current perspective on GCL and does not consider it a "burden" on working memory. However, for clarity, it continues to refer to cognitive efforts towards learning-relevant activities as GCL.

Theoretical Synthesis: Contextual Vertical Integration and Cognitive Load Profile

This study hypothesizes that CVI can optimize the cognitive load profile in clinical teaching. Revisiting basic science concepts in the direct context of their clinical application may manage ICL by activating preexisting knowledge. When existing schemas in long-term memory related to the current teaching material are activated, learners can "chunk" multiple elements of new information into a single element, thereby reducing the demands on limited working memory resources.[26,27]

CVI may also reduce ECL by addressing the split-attention effect–a source of ECL when learners must integrate information from temporally and/or spatially disparate sources.[28,29] By embedding scientific principles directly into clinical content, CVI could minimize this dissonance and foster efficient cognitive processing.

If CVI optimizes ICL and ECL, then the available working memory resources could be directed to support schema acquisition and enable meaningful learning, which are referred to in this study as GCL.

Objective

This "Contextual Vertical Integration in Clinical Education in Obstetrics and Gynecology" (CONVINCED-OG) trial, aims to evaluate the effects of CVI, compared to standalone clinical teaching without integration, on 3rd-year medical students' perceived intrinsic, extraneous, and germane cognitive loads during classroom-based teaching in undergraduate obstetrics and gynecology.

Methods

Study Setting

The CONVINCED-OG trial was conducted between February and May 2023 at two teaching units of Newcastle University, with both units delivering an identical curriculum and teaching methodology. To ensure instructional consistency, the lead author (M.I-K.) taught all study sessions at both units. Four Obstetrics and four Gynecology topics selected based on their perceived complexity (by previous cohorts) represented a core clinical topic from the undergraduate Obstetrics and Gynecology (O&G) curriculum defined by the Royal College of Obstetricians and Gynecologists[30] and the General Medical Council.[3]

Ethical Consideration

The Newcastle University Faculty of Medical Sciences Ethics Committee granted ethical approval for the study (Ref: 2465/28081/MEE8077). Students were given a participant information sheet and allowed to ask questions before obtaining written consent.

Participants

All third-year MBBS students assigned to O&G placements at both units during the study period were invited to participate. These students had completed the first two years of study, which focused on basic sciences taught through a case-based learning format.

During their third-year O&G placement, the students participated in structured, classroom-based teaching sessions covering core clinical topics delivered by case-based learning within a classroom setting, alongside clinical shadowing and practical experience. All 114 eligible students provided written informed consent. Students were eligible if they: 1)

 https://doi.org/10.2147/AMEP.S557502

were in the third year of study, 2) were enrolled at Newcastle University and 3) were attending O&G placements during the study period. Students from other universities or those with incomplete attendance during the study teaching sessions were excluded. Ultimately, 103 students with complete data were included in the final analysis.

Study Design

The study followed a cluster-randomized, topic-level crossover design (Figure 1 and Table 1). Each clinical topic represented an independent randomized controlled trial (RCT), hereafter referred to as "RCT-1" to "RCT-8", corresponding respectively to the four obstetric and four gynecological topics (Table 1). A computer-generated 3×8 Latin-square schedule allocated two cohorts from each unit to the integrated CVI format and one cohort to the standard, non-integrated format for Topic 1, then rotated the pattern for every topic, to ensure that (i) every cohort received four integrated and four standard sessions with an equal divide between Obstetrics and Gynecology, and (ii) every topic was delivered equally in both formats across the student cohorts.

The randomization list was prepared by a statistician not otherwise involved in the study, and the students remained blinded to the randomization status for each topic. The teacher learned the format only at the start of each session. In the non-integrated sessions, the clinical content was presented without integrating basic sciences. To ensure ethical parity, the students in the control group received the integrated version of each topic after the outcome measurement.

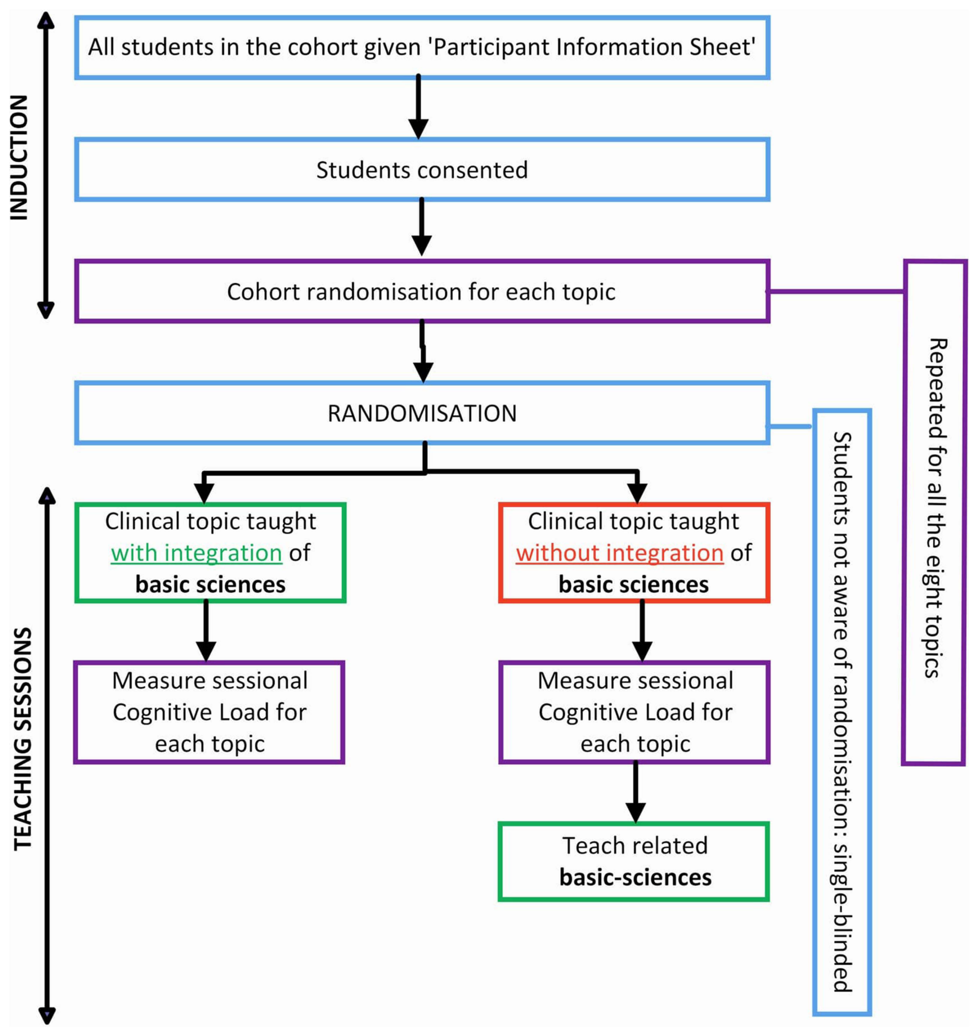

Figure 1 Study methodology.

Table 1 Cohort Randomization Characteristics for Each Clinical Topic

Randomized Controlled Trial (RCT)	Clinical Topic	Local Education Provider-Cohort (Number of Students) N - Total Number of Students Randomized to the Group	
		Taught with CVI	Taught without CVI
Obstetrics			
1.	Multifetal Gestation	X c1 (27), X-c3 (27), Y-c2 (7)	X-c2 (28), Y-c1 (6), Y-c3 (8)
		N=61	N=42
2.	Hypertensive Disorders in Pregnancy	X-c2 (28), Y-c1 (6), Y-c3 (8)	X-c1 (27), X-c3 (27), Y-c2 (7)
		N=42	N=61
3.	Postpartum Hemorrhage	X-c1 (27), X-c3 (27), Y-c2 (7)	X-c2 (28), Y-c1 (6), Y-c3 (8)
		N= 61	N=42
4.	Labour and Delivery	X-c2 (28), Y-c1 (6), Y-c3 (8)	X-c1 (27), X-c3 (27), Y-c2 (7)
		N=42	N=61
Gynaecology			
5.	Uterovaginal Prolapse & Urinary Incontinence	X-c2 (28), Y-c1 (6), Y-c3 (8)	X-c1 (27), X-c3 (27), Y-c2 (7)
		N=42	N=61
6.	Polycystic Ovarian Syndrome	X-c1 (27), X-c3 (27), Y-c2 (7)	X-c2 (28), Y-c1 (6), Y-c3 (8)
		N= 61	N=42
7.	Pelvic Inflammatory Disease	X-c2 (28), Y-c1 (6), Y-c3 (8)	X-c1 (27), X-c3 (27), Y-c2 (7)
		N=42	N=61
8.	Cervical Screening and Colposcopy	X-c1 (27), X-c3 (27), Y-c2 (7)	X-c2 (28), Y-c1 (6), Y-c3 (8)
		N= 61	N=42

Notes: X and Y are two different teaching hospitals. c1, c2, and c3 are the first, second, and third cohorts of third-year medical students. For example, X-c1 is the first cohort of medical students at teaching hospital X.

Intervention Design

Each of the eight topics was delivered as a single case-based learning session in which the basic-science explanation was woven directly into the unfolding case rather than presented as a separate mini-lecture. Development followed a streamlined Analyse, Design, Develop, Implement, Evaluate (ADDIE) cycle before the study commenced.[31] A gap analysis with focus groups of previous-year students and clinical teachers pinpointed aspects of each clinical condition that were frequently misunderstood; these informed a storyboard for each case that flagged "integration opportunities" for bringing in basic science mechanisms for each element of the clinical topic. A one-concept slide limited to a single diagram or pathway was designed for every opportunity according to Mayer's multimedia principles to minimize extraneous load.[32]

The full package (case script, lesson plan and slides) underwent staggered Delphi reviews, two topics at a time, with six subject-matter experts and six previous cohort students. Items were rated for relevance and clarity on a 5-point Likert scale; consensus was predefined as median ≥ 4 and inter-quartile range ≤ 1. All 56 items across the eight topics met this threshold by the second round, giving an overall content-validity index of 0.91. After minor wording tweaks, the validated materials were uploaded to SharePoint and used unchanged at both teaching units, with a single tutor (M. I-K.) delivering every session. Examples include pelvic-floor anatomy integrated during the examination phase of the

 https://doi.org/10.2147/AMEP.S557502

Table 2 Mapped Basic Sciences to Clinical Topics

Topic	Physiology*	Anatomy	Pharmacology	Bio-Chemistry	Pathology	Microbiology
Obstetrics						
Multifetal Gestation	✓	✓				
Hypertensive Disorders in Pregnancy	✓		✓	✓	✓	
Postpartum Hemorrhage	✓	✓	✓			
Labour & Delivery	✓	✓			✓	
Gynaecology						
Uterovaginal Prolapse & Urinary Incontinence	✓	✓	✓			✓
Polycystic Ovarian Syndrome	✓	✓	✓	✓		
Pelvic Inflammatory Disease		✓	✓			✓
Cervical Screening and Colposcopy		✓		✓	✓	✓

Note: *Physiology and Pathophysiology.

uterovaginal-prolapse case, and endocrine physiology, ovarian pathology and pharmacology woven throughout the polycystic-ovary-syndrome case. A complete mapping of basic-science domains to clinical topics appears in Table 2.

Minimizing Confounders and Bias

To minimize variability due to different instructors, the same instructor delivered all the sessions using standardized lesson plans to maintain consistency. Sessions were scheduled during the first two weeks of the six-week placement block to reduce the influence of accumulated clinical experience and to avoid bias associated with assessment stress in the last week.

Sessions included structured breaks to maintain engagement, and all multimedia materials were developed in accordance with Mayer's principles of multimedia learning.[32] These design controls aimed to minimize ECL from non-experimental sources, ensuring that any differences in perceived cognitive load could be attributed to the intervention.

Outcome Measures

Immediately after each session, students completed the self-reported validated eight-item Klepsch Cognitive Load Questionnaire[33] on paper, under supervision, to ensure standardized instructions and complete responses. The questionnaire was completed anonymously, with students providing only a self-generated identification code consisting of the middle four digits of their mobile number. This self reported instrument yielded separate scores for ICL, ECL and GCL.

Each item used a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree). Example items include "The session was very complex" (ICL), "The presentation of the learning content was very ineffective" (ECL), and "I tried to understand the relationships between the concepts" (GCL).[33] Subscale scores were calculated by summing the corresponding item responses (see Table 3).

In accordance with the instrument guide for preserving psychometric integrity, the students were not informed of the psychological constructs being measured.[33] This naïve reporting has been shown to yield strong internal consistency, with Cronbach's α value ranging from 0.70 to 0.86 across subscales, and strong construct validity (root-mean squared error of approximation of 0.02).[33] The utility of the instrument is further highlighted by its proven prognostic validity in capturing cognitive load variations attributable to changes in instructional design.[34]

While the interpretation of the GCL has evolved since the publication of Klepsch's questionnaire in 2017, from a distinct category of load to a conceptualization of schema-related processing or learning effort in 2019,[16,19] no updated

Table 3 Interpretation of Klepsch's Cognitive Load Questionnaire

Psychological Construct	No. of Items	Likert Scale	Score Range	Interpretation
Intrinsic Cognitive Load	2	1 to 7	2–14	Higher intrinsic load – impedes learning
Extraneous Cognitive Load	3	1 to 7	3–21	Higher extraneous load – impedes learning
Germane Cognitive Load	3	1 to 7	3–21	Higher germane load – facilitates learning

instruments have been developed or validated since. Therefore, this study retains the terminology of GCL to remain consistent with the instrument's structure and established psychometric properties, while the interpretation of GCL aligns with contemporary theory, representing the working memory resources allocated to learning-relevant processes.

Data Analysis

Sample size calculations for each intended statistical test were conducted using SAS version 9.4. The final sample of 103 students exceeded the minimum requirement to ensure 80% statistical power at a 95% confidence level. Normality of data distribution was assessed using the Kolmogorov–Smirnov test via the Social Science Statistics online platform.

Data analysis was conducted using SPSS version 29. Between-group comparisons of cognitive load scores (CVI vs non-CVI) for each RCT were performed using unpaired t-tests. When Levene's test indicated a violation of the assumption of homogeneity of variance, Welch's correction was applied. Within-subject differences in cognitive load perceptions—comparing integrated and non-integrated teaching sessions—were assessed using paired t-tests. The null hypothesis for each comparison stated that there would be no statistically significant difference in perceived cognitive load between students receiving clinical instruction with or without CVI. A p-value of < 0.05 was considered statistically significant.

The risk of Type I error arising from multiple topic-level comparisons was considered. To address this, convergent trends across topics were examined, and a pooled within-subject analysis was performed. As the study was designed to explore consistent patterns rather than to test independent hypotheses, formal multiple-comparison corrections were not applied.

To confirm the robustness of the results, corresponding non-parametric tests were performed: the Mann–Whitney *U*-test for between-group comparisons and the Wilcoxon signed-rank test for within-subject comparisons.

Results

Differences in Cognitive Load Between Integrated and Non-Integrated Teaching

Across eight topics, considered as separate RCTs, a consistent pattern was observed. Compared with those receiving non-integrated teaching, students taught with CVI reported lower perception of ICL and ECL and higher perceived GCL. The following sections detail these differences by cognitive load type.

Intrinsic Cognitive Load

Comparison of perceived ICL scores between integrated and non-integrated cohorts across the eight RCTs is presented in Table 4. For seven out of eight topics, students taught with CVI reported significantly lower ICL. The exception was the topic of "hypertensive disorders of pregnancy" (RCT-2), where the difference did not reach statistical significance ($p = 0.051$).

In RCT-2, RCT-7, and RCT-8, the data from the integrated groups did not follow a normal distribution. These data sets were analyzed using the Mann–Whitney *U*-test, which revealed statistically significant differences in ICL ($p < 0.001$) for RCT-7 and RCT-8 but not for RCT-2 ($p = 0.071$, $Z = -1.806$, Mann–Whitney $U = 1016$).

The largest reductions in ICL were observed in "pelvic inflammatory disease" (RCT-7, mean difference = –2.10, $p < 0.001$) and "cervical screening and colposcopy" (RCT-8, mean difference = –2.67, $p < 0.001$). These correspond to a 25%

 https://doi.org/10.2147/AMEP.S557502

Table 4 Differences in ICL Between Students Taught with and without CVI

Randomized Controlled Trial (RCT)	Topic	Mean ± Standard Deviation (for ICL) in Integrated Group (N -Sample Size)	Mean ± Standard Deviation (for ICL) in Non-Integrated Group (N -Sample Size)	t-value	p-value [Two-Tailed]	95% Confidence Interval	Cohens d [95% CI]
Topic 1	Multifetal Gestation	9.08 ± 1.706 N = 61	11.07 ± 1.552 N = 42	−6.031	<0.001 (Significant)	−2.64 to −1.335	1.645 [−1.63, −0.78]
Topic 2	Hypertensive Disorder of Pregnancy#	9.40 ± 1.939 N = 42	10.16 ± 1.908 N = 61	−1.972	0.051 (Not Significant)	−1.523 to 0.005	1.920 [−0.791, 0.002]
Topic 3	Postpartum Hemorrhage*	7.84 ± 2.465 N = 61	8.86 ± 2.465 N = 42	−2.342	0.022 (Significant)	−1.892 to −0.151	2.028 [−0.901, −0.103]
Topic 4	Labour and Delivery	8.21±1.815 N = 42	9.59±1.716 N = 61	−3.905	<0.001 (Significant)	−2.075 to −0.677	1.757 [−1.189, −0.374]
Topic 5	Urogynecology	7.98 ± 1.600 N = 42	10.30 ± 1.801 N = 61	−6.714	<0.001 (Significant)	−3.004 to −1.634	1.722 [−1.778, −0.909]
Topic 6	Polycystic Ovarian Syndrome	9.30 ± 0.969 N = 61	11.05 ± 1.766 N = 42	−4.626	<0.001 (Significant)	−2.504 to −1.001	1.890 [−1.339, −0.512]
Topic 7	Pelvic Inflammatory Disease#,*	6.29±1.274 N = 42	8.39±1.735 N = 61	−7.105	<0.001 (Significant)	−2.696 to −1.519	1.564 [−1.779, −0.910]
Topic 8	Cervical Screening Program and Colposcopy#	7.97±1.505 N = 61	10.64±2.046 N = 42	−7.647	<0.001 (Significant)	−3.370 to −1.982	1.745 [−1.976, −1.084]

Notes: #Indicates deviation from normal distribution as assessed by the Kolmogorov–Smirnov test; unpaired *t*-test results were cross-validated using Mann–Whitney U non-parametric testing. *Indicates violation of homogeneity of variance as determined by Levene's test; Welch's correction was applied to the unpaired *t*-test.

reduction in ICL due to CVI (Figure 2). The smallest reduction was in RCT-2 (hypertensive disorders of pregnancy, mean difference = 0.76, p = 0.051), where a 7.5% reduction was not statistically significant (Figure 2).

Extraneous Cognitive Load

As shown in Table 5, the reported ECL scores were significantly lower in the CVI condition for all eight topics ($p < 0.01$ in all cases). In RCT-5 and RCT-7, data from integrated cohorts were not normally distributed, and a parallel analysis with the Mann–Whitney *U*-test confirmed a statistically significant difference ($P < 0.001$).

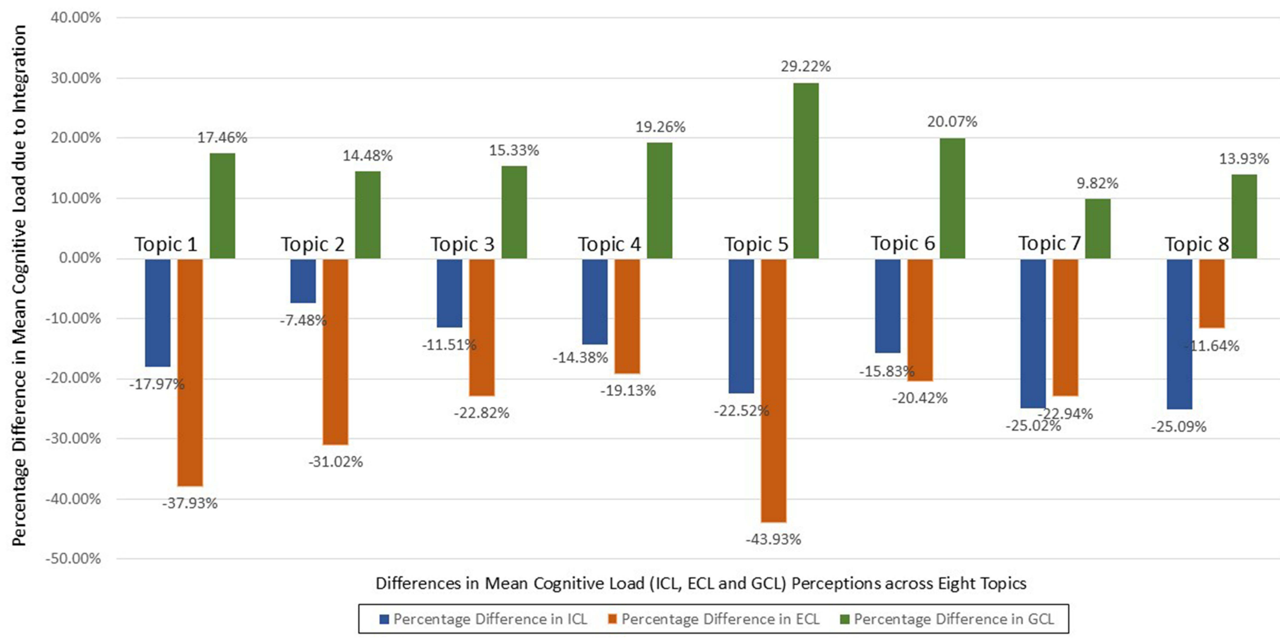

Figure 2 Percentage differences in intrinsic (ICL), extraneous (ECL), and germane cognitive load (GCL) between contextual vertical integration (CVI) and non-integrated teaching. CVI reduced ICL and ECL while increasing GCL across all eight topics, reflecting an optimized cognitive load profile.

The most substantial reductions were noted in the "urogynecology" session (RCT-5, mean difference = −5.54, $p < 0.001$), with a 43.9% reduction in ECL followed by a 37.9% reduction in multifetal gestation (RCT-1, mean difference = −4.67, $p < 0.001$). The smallest but still statistically significant reductions were observed in RCT-8 (cervical screening

Table 5 Differences in ECL Between Students Taught with and without CVI

Randomized Controlled Trial (RCT)	Topic	Mean ± Standard Deviation (of ECL) in Integrated Group (N -Sample Size)	Mean ± Standard Deviation (of ECL) in Non-Integrated Group (N -Sample Size)	t-value	p-value [Two-Tailed]	95% Confidence Interval	Cohens d [95% CI]
Topic 1	Multifetal Gestation	7.64 ± 2.727	12.31 ± 2.727	−8.541	<0.001 (Significant)	−5.755 to −3.586	2.727
		N = 61	N = 42				[−1.757, −0.796]
Topic 2	Hypertensive Disorder of Pregnancy*	8.38±2.230	12.15±3.182	−7.062	<0.001 (Significant)	−4.825 to −2.709	2.835
		N = 42	N = 61				[−0.796, −0.893]
Topic 3	Postpartum Hemorrhage	6.56 ± 1.893	8.50 ± 1.991	−5.011	<0.001 (Significant)	−2.712 to −1.174	1.933
		N = 61	N = 42				[−1.419, −0.586]
Topic 4	Labour and Delivery	6.81±1.581	8.44±1.420	−5.475	<0.001 (Significant)	−2.225 to −1.041	1.488
		N = 42	N = 61				[−1.516, −0.674]

(*Continued*)

Table 5 (Continued).

Randomized Controlled Trial (RCT)	Topic	Mean ± Standard Deviation (of ECL) in Integrated Group (N -Sample Size)	Mean ± Standard Deviation (of ECL) in Non-Integrated Group (N -Sample Size)	t-value	p-value [Two-Tailed]	95% Confidence Interval	Cohens d [95% CI]
Topic 5	Urogynecology[#]	7.07± 1.905	12.61±2.084	−13.712	<0.001 (Significant)	−6.336 to −4.734	2.013
		N = 42	N = 61				[−3.291, −2.200]
Topic 6	Polycystic Ovarian Syndrome	10.21 ± 2.984	12.83 ± 2.429	−4.714	<0.001 (Significant)	−3.723 to −1.518	2.772
		N = 61	N = 42				[−1.357, −0.529]
Topic 7	Pelvic Inflammatory Disease[#]	6.48±1.770	8.41±1.419	−6.139	<0.001 (Significant)	−2.558 to −1.309	1.571
		N = 42	N = 61				[−1.656, −0.800]
Topic 8	Cervical Screening Program and Colposcopy	11.30 ± 2.465	12.79 ± 2.893	−2.808	0.006 (Significant)	−2.544 to −0.438	2.647
		N = 61	N = 42				[−0.962, 0.161]

Notes: [#]Indicates deviation from normal distribution as assessed by the Kolmogorov–Smirnov test; unpaired *t*-test results were cross-validated using Mann–Whitney U non-parametric testing. *Indicates violation of homogeneity of variance as determined by Levene's test; Welch's correction was applied to the unpaired *t*-test.

and colposcopy, mean difference = −1.49, 11%), RCT-4 (labour and delivery, mean difference = −1.63, 19%) and RCT-6 (polycystic ovarian syndrome, mean difference = −2.62, 20%). Percentage differences are illustrated in Figure 2.

Germane Cognitive Load/ Learning Effort

Students exposed to CVI consistently reported higher GCL levels across all topics (Table 6), with statistically significant differences in each case ($p < 0.001$). The data from integrated cohorts in RCT-4, RCT-6, and RCT-8 and the non-integrated cohort in RCT-5 violated normal distribution. Hence, a parallel Mann–Whitney *U*-test was performed, which confirmed the statistically significant differences in GCL ($p < 0.001$) noted with unpaired t-tests in the same direction.

The most significant increase was observed in "urogynecology" (mean difference = +3.98, $p < 0.001$) and "labour and delivery" (mean difference = +2.85, $p < 0.001$) sessions, with 30% and 20% increases in GCL with CVI, respectively (Figure 2). The lowest increase of 7%, which was still significant, was observed in RCT-7 (mean difference = +1.61, $p<0.001$).

Pooled Within-Subject Analysis for Differences in Cognitive Load Between Integrated and Non-Integrated Teaching Conditions

To examine overall cognitive load differences, within-subject comparisons were conducted for all 103 students who experienced both integrated and non-integrated sessions. The Results are presented in Table 7.

A significantly lower ICL was noted when students were taught with integration (8.31 ± 1.44) than they were not (9.93 ± 1.56), $p < 0.001$. Similarly, students perceived lower ECL in sessions where the instructional design had CVI than in sessions without CVI (8.21 ± 2.14 vs 10.89 ± 2.01), $p < 0.001$.

In contrast, the average perceived GCL was significantly higher in the CVI teaching condition (M = 17.01 ± 1.78) than in the non-integrated teaching condition (M = 14.63 ± 1.44), $p < 0.001$.

The results of the paired *t*-test are summarized in Table 7. Parallel non-parametric testing by the Wilcoxon Signed Rank Test revealed significantly lower median ICL (Median =8.25 vs 10, Z = 8.258, $p<0.001$) and ECL (Median = 8 vs

 https://doi.org/10.2147/AMEP.S557502

Table 6 Differences in GCL Between Students Taught with and without CVI

Randomized Controlled Trial (RCT)	Topic	Mean ± Standard Deviation (of GCL) in Integrated Group (N -Sample Size)	Mean ± Standard Deviation (of GCL) in Non-Integrated Group (N -Sample Size)	t-value	p-value [Two-Tailed]	95% Confidence Interval	Cohens d [95% Confidence Interval]
Topic 1	Multifetal Gestation	17.08 ± 2.193	14.45 ± 2.360	5.797	<0.001 (Significant)	1.730 to 3.529	2.262
		N = 61	N = 42				[0.735, 1.584]
Topic 2	Hypertensive Disorder of Pregnancy*	16.76±2.356	14.64±1.623	5.069	<0.001 (Significant)	1.287 to 2.958	1.954
		N = 42	N = 61				[0.663, 1.504]
Topic 3	Postpartum Hemorrhage	17.30 ± 2.092	15.00 ± 2.095	5.467	<0.001 (Significant)	1.462 to 3.128	2.094
		N = 61	N = 42				[0.673, 1.515]
Topic 4	Labour and Delivery#	17.64±1.737	14.79±1.473	8.984	<0.001 (Significant)	2.225 to 3.487	1.585
		N = 42	N = 61				[1.333, 2.263]
Topic 5	Urogynecology#	17.60±2.220	13.62±1.968	9.552	<0.001 (Significant)	3.147 to 4.797	2.074
		N = 42	N = 61				[1.439, 2.385]
Topic 6	Polycystic Ovarian Syndrome#	16.75±2.211	13.95±1.834	6.763	<0.001 (Significant)	1.980 to 3.624	2.066
		N = 61	N = 42				[0.918, 1.788]
Topic 7	Pelvic Inflammatory Disease	18.0 ± 1.767	16.39 ± 1.584	4.825	<0.001 (Significant)	0.946 to 2.267	1.661
		N = 42	N = 61				[0.550, 1.380]
Topic 8	Cervical Screening Program and Colposcopy#	15.62±1.809	13.71±1.470	5.668	<0.001 (Significant)	1.241 to 2.577	1.680
		N = 61	N = 42				[0.711, 1.557]

Notes: #Indicates deviation from normal distribution as assessed by the Kolmogorov–Smirnov test; unpaired *t*-test results were cross-validated using Mann–Whitney U non-parametric testing. *Indicates violation of homogeneity of variance as determined by Levene's test; Welch's correction was applied to the unpaired *t*-test.

10.75, Z = 8.674, p < 0.001) and significantly higher median GCL (Median = 16.75 vs 14.75, Z = −8.773, p <0.001) in sessions where students were taught by CVI vs without CVI.

Discussion

This study evaluated the effect of CVI on students' perceptions of cognitive load during clinical teaching in undergraduate O&G. Across eight RCTs and pooled within-subject comparisons, students consistently reported lower ICL and ECL, alongside a higher GCL, when taught with CVI than when taught without integration. These findings provide empirical support for integrating basic sciences into clinical teaching from a cognitivist perspective. To our knowledge, this is the first study to evaluate the impact of vertical integration on optimizing cognitive load in undergraduate medical education.

The observed reduction in ICL suggests that embedding basic science explanations within clinical topics reduces the perceived complexity of the content. According to CLT, learners with prior relevant knowledge experience lower ICL, as they are better able to assimilate information.[21,25] CVI likely activates relevant existing schemas for previously learnt basic sciences, allowing clinical information to be processed more efficiently. This may enable working memory to treat multiple related elements of information as a single unit–a process referred to as "chunking".[26,27] For instance,

https://doi.org/10.2147/AMEP.S557502

Table 7 Comparison of Cognitive Loads Perceived by Students in Integrated vs Non-Integrated Teaching Sessions

Type of Cognitive Load	Sample Size (N)	Mean ± Standard Deviation of Individual Cognitive Load with Integration	Mean ± Standard Deviation of Individual Cognitive Load without Integration	t-value	p-value [Two-Tailed] (Significance of Difference)	95% Confidence Interval	Cohens d (95% Confidence Interval)
Intrinsic Cognitive Load (ICL) – maximum 14	103 pairs	8.31 ± 1.44	9.93 ± 1.56	−14.406	<0.001 (Significant)	−1.847 to −1.400	1.143 [−1.692, −1.144]
Extraneous Cognitive Load (ECL) – maximum 21	103 pairs	8.21 ± 2.14	10.89 ± 2.01	−14.700	<0.001 (Significant)	−3.038 to −2.315	1.848 [−1.724, −1.170]
Germane Cognitive Load (GCL) – maximum 21	103 pairs	17.01 ± 1.78	14.63 ± 1.44	21.688	<0.001 (Significant)	2.176 to 2.614	1.121 [1.784, 2.486]

integrating the pharmacology of oxytocin with its clinical use in managing postpartum hemorrhage allows learners to understand the rationale for its administration as a uterotonic, rather than memorizing isolated clinical facts such as dose, route, and adverse effects. In this way, CVI enables basic sciences to serve as a "cognitive glue," facilitating the construction of coherent and meaningful memory structures from otherwise disparate clinical facts.

CLT also posits that the interactivity of informational elements within a topic determines its ICL, with highly interactive topics being perceived as more complex.[26] Polycystic ovarian syndrome is one such condition that requires the integration of multiple physiological, biochemical, endocrine, and pharmacological concepts. CVI facilitates the sequential presentation of these basic science principles in relation to specific aspects of the condition. This allows learners to develop a structured understanding before synthesizing them into a comprehensive clinical script. In contrast, non-integrated teaching risks presenting all clinical elements simultaneously, which may overwhelm working memory. This benefit of CVI is consistent with the "isolated elements effect" described by Pollock,[35] where presenting information in an incrementally complex, sequenced manner reduces ICL and improves learning outcomes.[21,24]

The consistent reduction in ECL across all topics highlights the value of CVI-informed instructional design. We propose two mechanisms by which CVI may have reduced sources of ECL in this study. First, CVI likely mitigates the "redundancy effect", which occurs when learners must integrate separate sources of information that could be more efficiently understood if presented together.[21,36] For example, correlating early embryological development (basic sciences) with the classification of twins (clinical knowledge), which is usually taught in different years of medical school, imposes ECL on working memory. When presented together through CVI, however, learners can form a unified understanding of the embryology of chorionicity, amnionicity, and their clinical implications. Second, CVI appears to reduce the "split attention effect", which arises when interdependent content is taught in isolation, forcing learners to divide attention across sources.[21,28,29] For instance, understanding cervical dysplasia requires synthesizing the role of human papillomavirus (microbiology), cytological and histological changes (pathology), and colposcopic management (clinical component). CVI brings these elements together into a unified instructional narrative, reducing ECL and enhancing comprehension.

The observed increases in GCL further support the cognitive benefit of CVI. A higher GCL reflects allocation of working memory resources to learning-relevant processes and schema construction.[18,21,26] When interpreted alongside the reduction in loads on working memory (ie ICL and ECL), this suggests that CVI helps optimize the cognitive load profile, allowing information to be more effectively organized and consolidated in long-term memory.

There is strong evidence that clinicians draw on basic science knowledge when managing clinical conditions.[10,11,17,37–39] Schema theory suggests that linking new information to relevant prior knowledge (existing schemas) supports the formation of a "higher-order schema", which is a superior, interconnected structure of knowledge.[40] By reducing extraneous demands, CVI can potentially facilitate the development of a higher-order schema that encapsulates basic sciences and clinical knowledge. In contrast, a lower GCL in non-integrated instruction may result in multiple "lower-order schemas" for a single clinical condition.

Clinical competence depends on retrieving information from long-term memory and then applying it, accessing it from a unified "higher-order schema" is more cognitively efficient than retrieving it from several disjointed schemas.[41] Studies by Bassin and Lubitz,[42] Lisk et al[10] and Baghdady et al[15] support this by providing evidence of improved retention and application when teaching includes vertical integration. The consistent reduction in ICL and ECL, coupled with increased GCL observed in this study, aligns with these theoretical and empirical models, suggesting that CVI may support the formation of such higher-order schemas and thereby promote durable and transferable learning.

Unlike many previous studies that have evaluated vertical integration primarily through learning outcomes (ie the *product* of learning), this study investigates the *process* of learning by examining students' perceptions of cognitive load, an area that remains underexplored in the literature. The findings provide empirical support for integrating basic sciences into clinical teaching from a cognitivist perspective. To our knowledge, this is the first study to evaluate the impact of vertical integration on optimizing cognitive load in undergraduate clinical teaching. From a theoretical standpoint, it also extends the application of CLT to a real-life undergraduate clinical education setting, reinforcing its ecological validity and demonstrating its practical relevance for instructional design in medical education contexts.

Building on the significance of the findings of this study, their practical implications for curriculum design merit consideration. The principles of CVI can be operationalized through structured faculty development and deliberate collaboration between basic science and clinical educators.[6] Faculty training programs should focus on identifying "integration opportunities" within existing case-based or traditional clinical teaching sessions and on developing explanatory narratives that link underlying scientific mechanisms to clinical manifestations.

Cross-disciplinary curriculum planning meetings can ensure alignment of learning objectives and sequencing across pre-clinical and clinical years. Harden's Integration Ladder[12] provides a guiding framework for situating and progressively advancing the level of integration within programs. Established instructional-design models such as ADDIE[31] can guide the systematic development of integrated sessions, while Mayer's multimedia principles[32] can inform the creation of learning materials that minimize ECL. Embedding these processes within institutional faculty-development strategies could promote sustainable adoption of CVI across disciplines.

At this point, it is worth considering that some studies have reported no significant educational benefit from vertical integration.[43–45] These variations underscore the importance of *how* integration is implemented. The cognitive benefits demonstrated in this study should not be attributed to vertical integration per se, but rather to the degree of contextuality embedded within it. This is elucidated by Cheung et al[16] who demonstrated improved learning outcomes when basic sciences were taught concurrently with clinical content, as opposed to being delivered beforehand. The distinction between basic sciences and clinical knowledge being delivered in "temporal proximity" and those that are "intentionally interwoven" is crucial to achieving the cognitive benefits of integration.

This study has several limitations. It was conducted within a single institution, focused on one specialty, and involved a single-year cohort, which could limit generalizability. Additionally, the findings were based on classroom-based teaching, which may not directly translate to less structured workplace-based learning. The outcomes measured were perceptual rather than objective, and self-reported data may be influenced by social desirability or recall bias. Although instructional delivery was standardized, learner engagement, prior knowledge, and preferred learning styles may have introduced variability. Having a single instructor was aimed at minimizing variability in teaching quality, however, this aspect of the design may have introduced a potential instructor effect. In addition, participants' awareness of being part of a research study could have elicited a Hawthorne effect, potentially influencing engagement or self-reported perceptions.

 https://doi.org/10.2147/AMEP.S557502

Moreover, this study focused only on knowledge acquisition - a single facet of the broader competency framework required in medical education, including communication, procedural skills, and professionalism.

When conducting this research, Klepsch's instrument was widely accepted as a validated tool for measuring cognitive load subtypes.[34,46] Although concerns about its factorial structure have since been raised and newer instruments such as the Krieglstein scale have appeared,[47] the patterns of cognitive load differences noted in this study provide empirical evidence that Klepsch's instrument can discriminate meaningfully between instructional conditions.

Further research is needed to evaluate the applicability of CVI across other specialties, institutions, and health professions education. Mixed-methods designs may help capture learners' reflections and provide deeper insights into how CVI is experienced in diverse educational contexts. This study focused on short-term cognitive perceptions; therefore, future longitudinal research should explore whether the immediate optimisation of cognitive load profiles during teaching sessions leads to improved long-term knowledge retention and clinical performance. We also recognize the value of emerging instruments for cognitive load assessment and recommend that future multi-center replications incorporate these newer tools. Additionally, there remains a broader gap in literature linking perceived cognitive load to both immediate and sustained educational outcomes.

Conclusion

This CONVINCED-OG trial provides compelling evidence that CVI is an effective instructional strategy for optimizing cognitive load in undergraduate clinical education. By reducing the cognitive load associated with content complexity and suboptimal instructional design, CVI enables cognitive efforts to be redirected toward meaningful learning.

These findings underscore the need to move beyond merely co-locating basic science and clinical content toward deliberately designed, contextually integrated approaches that support cognitive efficiency. Curriculum developers and educators should consider embedding CVI principles, particularly in cognitively demanding topics. This aligns teaching with students' cognitive architecture to enhance learning outcomes. As medical education continues to evolve, CVI offers a practical, theory-informed framework to bridge basic sciences and clinical practice in a way that is pedagogically sound, cognitively optimized, and learner-centered.

Data Sharing Statement

The datasets generated and analyzed during the current study are not publicly available at this time, as they are being further analyzed for a separate publication addressing a different research question nested within the CONVINCED-OG Trial. However, the data are available from the corresponding author upon reasonable request.

Ethics Approval

Ethical approval for the study was granted by the Ethics Committee and Research Management Group of the Faculty of Medical Sciences, Newcastle University (Ref: 2465/28081/MEE8077). The study was performed in accordance with Newcastle University's guidelines and regulations.

Consent to Participate

All participants were informed of their involvement in the research and provided written informed consent. Data were anonymized at the point of collection, and explicit consent for publication was obtained from all participants.

Funding

No funding was received for this study.

Disclosure

The authors declare that they have no competing interests in this work.

References

1. Liaison Committee on Medical Education (LCME). Functions and structure of a medical school. 2025. Available from: https://lcme.org/publications/. Accessed April 30, 2025.
2. Australian Medical Council (AMC). Standards for assessment and accreditation of primary medical programs 2 standards for assessment and accreditation of primary medical programs standards for assessment and accreditation of primary medical programs.
3. General Medical Council (GMC). Outcomes for graduates 2018. UK: General Medical Council; 2018. Available from: https://www.gmc-uk.org/-/media/documents/dc11326-outcomes-for-graduates-2018_pdf-75040796.pdf. Accessed April 30, 2025.
4. Pusparajah P, Goh BH, Lee LH, et al. Integrating the basic and clinical sciences throughout the medical curriculum: contemplating the why, when and how. *Prog Drug Discov Biomed Sci*. 2022;5(1). doi:10.36877/PDDBS.A0000308
5. Finnerty EP, Chauvin S, Bonaminio G, Andrews M, Carroll RG, Pangaro LN. Flexner revisited: the role and value of the basic sciences in medical education. *Acad Med*. 2010;85(2):349–355. doi:10.1097/ACM.0B013E3181C88B09
6. Brauer DG, Ferguson KJ. The integrated curriculum in medical education: AMEE Guide No. 96. *Med Teach*. 2015;37(4):312–322. doi:10.3109/0142159X.2014.970998
7. Custers EJFM. Long-term retention of basic science knowledge: a review study. *Adv Health Sci Educ*. 2010;15(1):109–128. doi:10.1007/S10459-008-9101-Y
8. Ling Y, Swanson DB, Holtzman K, Bucak SD. Retention of basic science information by senior medical students. *Acad Med*. 2008;83(10 Suppl):S82–5. doi:10.1097/ACM.0B013E318183E2FC
9. D'Eon MF. Knowledge loss of medical students on first year basic science courses at the University of Saskatchewan. *BMC Med Educ*. 2006;6(1):1–6. doi:10.1186/1472-6920-6-5/FIGURES/1
10. Lisk K, Agur AMR, Woods NN. Exploring cognitive integration of basic science and its effect on diagnostic reasoning in novices. *Perspect Med Educ*. 2016;5(3):147–153. doi:10.1007/S40037-016-0268-2
11. Woods NN, Brooks LR, Norman GR. It all make sense: biomedical knowledge, causal connections and memory in the novice diagnostician. *Adv Health Sci Educ*. 2007;12(4):405–415. doi:10.1007/S10459-006-9055-X
12. Harden RM. The integration ladder: a tool for curriculum planning and evaluation. *Med Educ*. 2000;34(7):551–557. doi:10.1046/J.1365-2923.2000.00697.X
13. Davis MH, Harden RM. Planning and implementing an undergraduate medical curriculum: the lessons learned. *Med Teach*. 2003;25(6):596–608. doi:10.1080/0142159032000144383
14. Woods NN. Science is fundamental: the role of biomedical knowledge in clinical reasoning. *Med Educ*. 2007;41(12):1173–1177. doi:10.1111/J.1365-2923.2007.02911.X
15. Baghdady MT, Carnahan H, Lam EWN, Woods NN. Integration of basic sciences and clinical sciences in oral radiology education for dental students. *J Dent Educ*. 2013;77(6):757–763. doi:10.1002/J.0022-0337.2013.77.6.TB05527.X
16. Cheung JJH, Kulasegaram KM, Woods NN, Brydges R. Why content and cognition matter: integrating conceptual knowledge to support simulation-based procedural skills transfer. *J Gen Intern Med*. 2019;34(6):969–977. doi:10.1007/S11606-019-04959-Y
17. Woods NN, Neville AJ, Levinson AJ, Howey EHA, Oczkowski WJ, Norman GR. The value of basic science in clinical diagnosis. *Acad Med*. 2006;81(10 SUPPL):S124–7. doi:10.1097/00001888-200610001-00031
18. Sweller J. The development of cognitive load theory: replication crises and incorporation of other theories can lead to theory expansion. *Educ Psychol Rev*. 2023;35(4):1–20. doi:10.1007/S10648-023-09817-2/TABLES/1
19. Sweller J. Cognitive load during problem solving: effects on learning. *Cogn Sci*. 1988;12(2):257–285. doi:10.1207/S15516709COG1202_4
20. Piaget J. Readings on the development of children. In: Gauvain M, Cole M, editors. *Readings on the Development of Children*. 2nd ed. W. H. Freeman and Company; 1964:26–28.
21. Sweller J, van Merriënboer JJG, Paas F. Cognitive architecture and instructional design: 20 years later. *Educ Psychol Rev*. 2019;31(2):261–292. doi:10.1007/S10648-019-09465-5/FIGURES/1
22. Atkinson RC, Shiffrin RM. Human memory: a proposed system and its control processes. *Psychol Learn Motiv Adv Res Theory*. 1968;2:89–195. doi:10.1016/S0079-7421(08)60422-3
23. Sweller J, Van Merrienboer JJG, Paas FGWC. Cognitive architecture and instructional design. *Educ Psychol Rev*. 1998;10(3):251–296. doi:10.1023/A:1022193728205
24. Van Merriënboer JJG, Sweller J. Cognitive load theory in health professional education: design principles and strategies. *Med Educ*. 2010;44(1):85–93. doi:10.1111/J.1365-2923.2009.03498.X
25. Sweller J. Cognitive load theory: recent theoretical advances. In: Plass J, Moreno R, Brunken R, editors. *Cognitive Load Theory*. Vol. 9780521860239. Cambridge University Press; 2010:29–47. doi:10.1017/CBO9780511844744.004
26. Kalyuga S. Schema acquisition and sources of cognitive load. In: Plass J, Moreno R, Brünken R, editors. *Cognitive Load Theory*. Cambridge University Press; 2010:48–64. doi:10.1017/CBO9780511844744.005
27. Ericsson KA, Kintsch W. Long-term working memory. *Psychol Rev*. 1995;102(2):211–245. doi:10.1037/0033-295X.102.2.211
28. Pouw W, Rop G, de Koning B, Paas F. The cognitive basis for the split-attention effect. *J Exp Psychol Gen*. 2019;148(11):2058–2075. doi:10.1037/XGE0000578
29. Ayres P, Sweller J. The split-attention principle in multimedia learning. In: Mayer R, editor. *The Cambridge Handbook of Multimedia Learning*. 2nd ed. Cambridge University Press; 2014:206–226. doi:10.1017/CBO9781139547369.011
30. Royal College of Obstetricians and Gynaecologists. Undergraduate curriculum. UK: Royal College of Obstetricians and Gynaecologists; 2009. Available from: https://www.rcog.org.uk/media/xvrnmpqj/rcog-undergraduate-curriculum-final-version-2021.pdf. Accessed April 30, 2025.
31. Molenda M. In search of the elusive ADDIE model. *Perform Improv*. 2003;42(5):34–36. doi:10.1002/PFI.4930420508
32. Mayer RE, Fiorella L. Principles for reducing extraneous processing in multimedia learning: coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In: Mayer R, editor. *The Cambridge Handbook of Multimedia Learning*. 2nd ed. Cambridge University Press; 2014:279–315. doi:10.1017/CBO9781139547369.015
33. Klepsch M, Schmitz F, Seufert T. Development and validation of two instruments measuring intrinsic, extraneous, and germane cognitive load. *Front Psychol*. 2017;8:294028. doi:10.3389/FPSYG.2017.01997/BIBTEX

34. Klepsch M, Seufert T. Understanding instructional design effects by differentiated measurement of intrinsic, extraneous, and germane cognitive load. *Instr Sci*. 2020;48(1):45–77. doi:10.1007/S11251-020-09502-9/TABLES/11
35. Pollock E, Chandler P, Sweller J. Assimilating complex information. *Learn Instr*. 2002;12(1):61–86. doi:10.1016/S0959-4752(01)00016-0
36. Chandler P, Sweller J. Cognitive Load Theory and the Format of Instruction. *Cogn Instr*. 1991;8(4):293–332. doi:10.1207/S1532690XCI0804_2
37. Danielson JA, Burzette RG, Bailey MR, et al. Basic science knowledge underlies clinical science knowledge and clinical problem solving: evidence from veterinary medicine. *Adv Health Sci Educ*. 2024;30(1):151–169. doi:10.1007/S10459-024-10334-2
38. Alele FO, Albert FA, Anderson E, et al. The perceived relevance, utility and retention of basic sciences in general practice. *BMC Med Educ*. 2024;24(1):1–12. doi:10.1186/S12909-024-05750-2/TABLES/3
39. Woods NN, Brooks LR, Norman GR. The role of biomedical knowledge in diagnosis of difficult clinical cases. *Adv Health Sci Educ*. 2007;12(4):417–426. doi:10.1007/S10459-006-9054-Y
40. Schmidt HG, Norman GR, Boshuizen HPA. A cognitive perspective on medical expertise: theory and implications. *Acad Med*. 1990;65(10):611–621. doi:10.1097/00001888-199010000-00001
41. Schmidt HG, Boshuizen HPA. Encapsulation of Biomedical Knowledge. In: Evans DA, Patel VL, editors. *Advanced Models of Cognition for Medical Training and Practice*. Springer; 1992:265–282. doi:10.1007/978-3-662-02833-9_15
42. Bassin SR, Lubitz S. Integrating basic science into endocrine clinical electives. *Med Sci Educ*. 2023;33(4):835. doi:10.1007/S40670-023-01824-W
43. Amini M, Mikaeili F, Handjani F, Hatam G, Asgari Q. The effect of integration of basic and clinical aspects of a specific topic in a parasitology course on medical students learning: a randomized controlled trial. *J Educ Health Promot*. 2021;10(1):390. doi:10.4103/JEHP.JEHP_1656_20
44. Brooks WS, Woodley KTCP, Jackson JR, Hoesley CJ. Integration of gross anatomy in an organ system-based medical curriculum: strategies and challenges. *Anat Sci Educ*. 2015;8(3):266–274. doi:10.1002/ASE.1483
45. Kercheval JB, Mott NM, Kim EK, et al. Students' perspectives on basic and clinical science integration when step 1 is administered after the core clerkships. *Teach Learn Med*. 2023;35(2):117–127. doi:10.1080/10401334.2022.2030235
46. Kastaun M, Meier M, Küchemann S, Kuhn J. Validation of cognitive load during inquiry-based learning with multimedia scaffolds using subjective measurement and eye movements. *Front Psychol*. 2021;12:703857. doi:10.3389/FPSYG.2021.703857/BIBTEX
47. Krieglstein F, Beege M, Rey GD, Sanchez-Stockhammer C, Schneider S. Development and validation of a theory-based questionnaire to measure different types of cognitive load. *Educ Psychol Rev*. 2023;35(1):1–37. doi:10.1007/S10648-023-09738-0