Enigma in Education
https://enigma.or.id/index.php/edu

e-ISSN: 3026-5827

ORIGINAL ARTICLE
Enigma in Education | Vol. 4, No. 2 (2026) | pp. 69–86

Instructional Alignment of Learning Objectives, Practice Tasks, and Assessments in an Open-Access Solid-State Chemistry Course: A Two-Dimensional Revised Bloom's Taxonomy Analysis

Sarah Armalia[1*], Kevin Setiawan[2], Neva Dian Permana[3]

[1]Department of Natural Sciences, Barelang Study Center, Tanjung Pinang, Indonesia
[2]Department of Chemical Engineering, CMHC Research Center, Palembang, Indonesia
[3]Department of Social Sciences, CMHC Research Center, Palembang, Indonesia

ARTICLE INFO

Article history
Received: August 21, 2026
Revised: September 20, 2026
Accepted: October 14, 2026
Available online: October 28, 2026
Published: November 1, 2026

Keywords
Chemistry education; constructive alignment; open educational resources; procedural knowledge; Revised Bloom's Taxonomy

Corresponding author
Sarah Armalia
saraharmalia@gmail.com

ORCID
0009-0007-8307-6578

DOI
https://doi.org/10.61996/edu.v4i2.136

ABSTRACT

Background. Judging instructional alignment among learning objectives, practice tasks, and assessments requires attention to substantive task demands, and surface action-verb frequencies alone are not enough.

Objective. To evaluate instructional alignment across the written curriculum of MIT OpenCourseWare 3.091SC (Introduction to Solid-State Chemistry) with the two-dimensional Revised Bloom's Taxonomy.

Methods. A rubric-based documentary curriculum analysis covered 35 sessions, nine modules, 176 explicit learning objectives, and 62 task documents (213 pages). The documents held 260 parent problems and 494 labeled task units, from which five within-type duplicates, 25 figure-restricted units, and seven context-insufficient units were excluded, leaving 457 unique codable task units, including 345 primary tasks (253 homework, 33 quiz, and 59 module examination units), alongside 175 codable objectives. Taxonomic distributions were evaluated with 24-cell overlap, 5,000 module-cluster bootstrap resamples, leave-one-module-out jackknife bounds, and six sensitivity specifications.

Results. Learning objectives emphasized C2 Understand (95/175, 54.3%) and Conceptual knowledge (101/175, 57.7%), whereas primary tasks emphasized C3 Apply (237/345, 68.7%) and Procedural knowledge (250/345, 72.5%). Aggregating C1–C3 covered 330/345 primary tasks (95.7%), yet C1–C2 made up only 93/345 (27.0%). Mean module-level 24-cell overlap was 0.620 (95% bootstrap CI: 0.490–0.743) for objectives with practice and 0.595 (0.498–0.690) for objectives with assessments. The mean C3 difference between assessments and practice was +0.6 percentage points (95% CI: −10.6 to +13.0).

Conclusion. Written tasks operationalize conceptual objectives mainly through quantitative and representational C3 application. No systematic C3 escalation from homework to examinations was found. Dual-rater calibration is a planned next step.

Copyright: © 2026 The Author(s). Published by Enigma Institute on behalf of Enigma in Education. The authors retain copyright of their work.

Access license: This is an open access article distributed under the terms of the Creative Commons Attribution–NonCommercial–ShareAlike 4.0 International License (CC BY-NC-SA 4.0), which permits non-commercial use, sharing, adaptation and reproduction in any medium, provided the original author and source are credited and any derivative work is distributed under the same license.

1. Introduction

1.1 Constructive Alignment and Task Demand in Undergraduate Chemistry Education

Constructive alignment rests on three connected elements: intended learning outcomes that state what students should be able to do, learning activities that engage students in the cognitive operations those outcomes require, and assessment tasks that elicit evidence that students can perform them.[1,2] In undergraduate chemistry and materials science, achieving this fit takes more than matching topical keywords across syllabus headings and examination papers.[3–7] Consider a session objective that asks students to explain the physical mechanism of atomic bonding or of lattice diffraction. An examination item that only asks for numerical constants to be substituted into a supplied formula does not satisfy it.[5,6] The reverse also holds. In quantitative physical sciences such as solid-state chemistry, computational problem sets and representational exercises are often the main route. Through them, abstract models are put into practice and tested.[8–10] A divergence in taxonomic labels between a conceptual objective and a quantitative task therefore does not by itself show poor curriculum design. It shows how a discipline turns theoretical comprehension into procedural execution.[1,5,11,12]

Evaluating this translation in OpenCourseWare (OCW) and Open Educational Resources (OER) matters for science education worldwide.[13,14] Self-directed learners, secondary teachers, and university instructors adopt open course packages as benchmark syllabi, problem banks, and self-assessment repositories, and one example is Massachusetts Institute of Technology (MIT) OpenCourseWare 3.091SC (Introduction to Solid-State Chemistry).[8,9] Even with this wide reuse, most curriculum evaluations rely on qualitative impressions of a syllabus or on automated counts of isolated action verbs.[4–6] Classifying objectives and examination questions by matching opening verbs ('define', 'calculate', 'explain', or 'construct') against fixed verb

lists can misstate the actual cognitive demand of a task.[2–6] In solid-state chemistry, an item that begins with 'construct' or 'draw' may only require plotting a binary eutectic phase diagram from fixed invariant temperatures. It may equally ask for a standard Lewis electron-dot structure drawn by taught convention. Both are rule-governed application, not open-ended creative synthesis.[2–6]

The relation between conceptual understanding and procedural execution can be framed with Johnstone's triplet: macroscopic phenomena, submicroscopic atomic and lattice models, and symbolic-mathematical representations.[15–19] In introductory solid-state chemistry, students cannot observe electron orbitals, Bravais lattice planes, point defects, or glass network connectivity directly. They reach these constructs through symbolic equations, crystallographic indices, and phase diagrams.[8,9,11,12,15–18] Recent international studies of undergraduate chemistry assessment point the same way, including Bloom-blueprinted first-year general chemistry examinations in the United Kingdom, longitudinal studies of organic chemistry assessment, analyses of three-dimensional learning progressions, and chemical engineering reforms of constructive alignment, and instructors in these studies state course goals in qualitative, conceptual language, yet rely on quantitative problem solving and symbolic representation as the main medium of assessment.[5,6,20,21] Evaluating alignment in chemistry OER therefore needs a two-dimensional taxonomic lens. It must show whether procedural tasks are disconnected rote drills or deliberate quantitative operationalizations of conceptual models.[1–4,22]

1.2 Two-Dimensional Revised Bloom's Taxonomy and the Disaggregation of Lower-Order Thinking Skills

To move beyond verb matching, Anderson and Krathwohl's Revised Bloom's Taxonomy separates cognitive demand into two intersecting dimensions. The Cognitive Process Dimension has six categories: Remember (C1), Understand (C2), Apply (C3), Analyze (C4), Evaluate (C5), and Create (C6), and the Knowledge Dimension has four: Factual, Conceptual, Procedural, and Metacognitive.[2,3] Crossing them gives a 24-cell taxonomic matrix. The matrix separates knowing chemical terminology and constants (Factual knowledge) from comprehending interatomic models and phase relationships (Conceptual knowledge). It also sets both apart from executing stoichiometric algorithms or crystallographic conventions (Procedural knowledge) and from monitoring one's own problem-solving strategy (Metacognitive knowledge).[2,3,15–18] The six process categories form an ordinal typology of qualitative cognitive operations, not an equal-interval scale. Arithmetic means of category indices, for example treating C1 to C6 as integer scores 1.0 to 6.0, are psychometrically invalid.[2,3,23]

A common distortion in secondary and higher education curriculum research comes from merging C1 Remember, C2 Understand, and C3 Apply into one undifferentiated block called Lower-Order Thinking Skills (LOTS). C4 Analyze, C5 Evaluate, and C6 Create are then set apart as Higher-Order Thinking Skills (HOTS).[5,6,23] The merger is costly in introductory general and solid-state chemistry, where mastery of quantitative problem solving, unit conversion, defect equilibria, and crystallographic indexing is foundational. Placing C3 Apply with C1 Remember and C2 Understand hides the difference between rote recall and multi-step quantitative application.[5,6] Suppose 95% of course tasks fall within C1–C3. A coarse LOTS/HOTS dichotomy could then falsely suggest that the course demands only rote memorization. In fact nearly 70% of tasks require quantitative and representational C3 application.[5,6,22] Separating C3 Apply from C1–C2, while tracking the Conceptual-to-Procedural shift across the Knowledge Dimension, is therefore essential for an accurate analysis that does not overclaim.[2,3,22,24]

1.3 Study Scope, Research Questions, and Epistemic Boundaries

Advances in computational text analysis and rubric-constrained classification now allow systematic evaluation of educational objectives and assessment items across whole university courses.[25–32] Even so, documentary taxonomic coding depends on operational rubric rules, preserved problem context, and dominant-category criteria until independent domain experts cross-check it.[27–32] Open courseware also leaves out proprietary third-party textbook figures and live classroom discourse, so documentary analyses of such material must therefore limit their conclusions to the written instructional tasks.[13,14,33–39]

Guided by these principles, this study applies a rubric-based, two-dimensional Revised Bloom's Taxonomy analysis to the complete written instructional corpus of MIT OpenCourseWare 3.091SC (Introduction to Solid-State Chemistry).[8,9] It addresses four research questions. (1) How do the distributions of cognitive process (C1–C6) and knowledge dimension (Factual, Conceptual, Procedural, Metacognitive) compare across explicit session learning objectives, homework practice problems, module quizzes, module examinations, supplemental problems, and the cumulative final examination? (2) To what extent does separating C3 Apply from C1–C2 change the reading of LOTS dominance in introductory solid-state chemistry? (3) What is the module-level 24-cell distributional overlap among learning objectives, practice tasks, and summative assessments across the nine thematic modules, and is there statistically stable evidence that C3 Apply escalates from homework practice to assessments? (4) How sensitive are task cognitive tier proportions and module alignment estimates to boundary-ambiguity exclusions, wider corpus scope, within- and cross-type deduplication, parent-item aggregation, and leave-one-module-out perturbations? The study evaluates written task demands within the 3.091SC curriculum. It does not measure student cognitive load, empirical item difficulty, classroom pedagogy, or realized learning gains.

2. Methods

2.1 Curriculum Research Design, Modular Scope, and Copyright Compliance

This study is a systematic, rubric-based documentary curriculum analysis of a benchmark open-access undergraduate chemistry course. The course is MIT OpenCourseWare (OCW) 3.091SC, Introduction to Solid-State Chemistry, developed at the Massachusetts Institute of Technology.[8,9] Its curriculum is a self-contained instructional sequence. It combines explicit session-level learning objectives, formative homework problem sets, self-assessment module quizzes, module examinations, supplemental problem sets, and a cumulative final examination across fundamental solid-state and materials chemistry topics.[8,9] Analyzing the complete written curriculum allows a direct comparison of intended learning outcomes with the operational task demands of formative and summative assessments across an entire undergraduate physical science course.

The sampling frame was the complete census of the 3.091SC instructional syllabus. It covers 35 instructional sessions organized into nine thematic modules: (1) Atomic Structure, (2) Bonding and Molecules, (3) Electronic Materials, (4) Crystalline Materials, (5) Amorphous Materials, (6) Reactions and Kinetics, (7) Aqueous Solutions, (8) Organic Materials, and (9) Solid Solutions, plus the cumulative final examination.[8,9] Every instructional session and assessment module was included. Sessions 7 and 27, among the 35, served as examination or review sessions and have no separate homework problem documents. External textbook reading problems, lecture videos, worked solution keys, and external web links were not used to generate standalone task units, so the analytical denominator reflects written course problems. The census comprised 62 standalone

written task documents spanning 213 pages: 33 homework problem sets, 10 module quizzes (Module 8 has two separate quiz documents), nine module examinations, nine supplemental problem sets, and one cumulative final examination.

MIT OpenCourseWare materials carry a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) license, and third-party copyrighted figures and tables are explicitly excluded from it.[13,14] Open-access courseware raises two issues that closed institutional item banks do not: intellectual property boundaries and multimodal representation demands.[13,14,40] Scientific diagrams, crystal structure illustrations, and phase diagrams reproduced from third-party publications stay under attribution restrictions.[13,14] To respect intellectual property and avoid speculative coding of missing diagrams, 25 segmented task units whose interpretation depended on third-party attributed figures or restricted visual excerpts were withheld from primary taxonomic classification, and no third-party proprietary diagrams are reproduced here. In solid-state chemistry, a visual prompt often decides what a question asks. It may call for recalling a standard lattice type (C1), interpreting a band diagram (C2), plotting a tie-line composition with the lever rule (C3), or diagnosing diffraction peak absences (C4). The 25 third-party figure-restricted units and the seven context-insufficient units were withheld from primary classification. Their visual content was not guessed, and no low-level C1 codes were imputed. This prevents artificial distortion of taxonomic proportions.

2.2 Unit Segmentation, Deduplication Rules, and Context Preservation

An unambiguous analytical denominator needs explicit unit boundary rules, and these were applied separately to learning objectives and to written tasks. A learning objective unit is a single explicit statement listed under the learning objectives of a session syllabus. This gave 176 explicit objective units across the 35 sessions. A written task unit is a lettered sub-question (a through f) within a multi-part problem. When a problem has no lettered sub-questions, the whole parent question stem is one unit. Numbered or Roman-numeral answer options stay attached to their parent unit and are not separate units. When a lettered sub-question was segmented, the full text of its parent stem was kept as analytical context. Classification could then weigh the sub-question against all given physical constants, chemical equations, and scenario constraints. Segmenting the 62 written documents under these rules produced 260 parent problem stems and 494 initial labeled task units. The 494 task units are 494 sub-questions or unsplit parent prompts nested within 260 parent problems. They are not 494 statistically independent questions or student test responses.

Segmentation was verified across all 260 parent problems. The check confirmed that chemical state symbols, measurement units, and internal cross-references did not split continuous prompts. Each segmented unit was linked to its thematic module, its instructional stage (learning objective, homework practice, module quiz, module examination, supplemental problem set, or cumulative final examination), its parent problem context, a normalized text representation, its dominant Cognitive Process category, its dominant Knowledge Dimension category, a substantive coding rationale, and a boundary-ambiguity indicator. Within-type deduplication normalized letter case, whitespace, and punctuation across all task units of the same instructional stage. It identified and removed five exact text duplicates within the same task type. Semantic paraphrases, and isomorphic calculations with altered numerical parameters across different problem sets, were examined separately in the sensitivity analyses.

2.3 Two-Dimensional Revised Bloom's Taxonomy Rubric and Operational Coding Protocol

Classification used a two-dimensional operational rubric grounded in Anderson and Krathwohl's Revised Bloom's Taxonomy and in chemistry education assessment frameworks.[2–6] Each codable unit received one dominant Cognitive Process category (C1 Remember, C2 Understand, C3 Apply, C4 Analyze, C5 Evaluate, or C6 Create) and one dominant Knowledge Dimension category (Factual, Conceptual, Procedural, or Metacognitive). The process rubric set the following boundaries. C1 Remember was limited to retrieving specific chemical terminology, atomic constants, or definitions, with no calculation or model interpretation. Questions that begin with 'Name' or 'Identify' but require stoichiometric calculation or orbital configuration rules were not coded C1. C2 Understand covered explaining physical or chemical relationships, summarizing bonding mechanisms, or interpreting crystal or molecular representations without executing a quantitative algorithm. The surface verb 'explain' alone did not raise an item to higher-order thinking. C3 Apply covered executing quantitative equations, stoichiometric or thermodynamic algorithms, or taught representational conventions. Drawing standard Lewis structures, sketching unit-cell planes, or plotting binary phase diagrams from fixed data points is coded C3 Apply, not C6 Create.

Higher-order codes required verifiable analytical, evaluative, or generative demands beyond routine procedure.[2–6] C4 Analyze was assigned when a task required diagnosing patterns from experimental data, such as indexing X-ray diffraction reflection rules to determine Bravais lattice centering. It was also assigned for integrating multi-variable structure–property relationships when the solution pathway was not pre-specified. C5 Evaluate required judging a chemical model, a material selection, or an experimental claim against explicit criteria. The word 'justify' attached to a routine calculation or a single-concept explanation was not enough. C6 Create required designing an open-ended material processing sequence or a novel experimental solution under several physical constraints. On the Knowledge Dimension, Factual knowledge meant isolated terms, symbols, and constants. Conceptual knowledge meant classifications, principles, bonding theories, and structure–property models. Procedural knowledge meant quantitative calculation methods, algebraic derivations, and representational drawing conventions. Metacognitive knowledge required explicit reflection on learning strategies or self-monitoring of cognitive heuristics.

Taxonomic coding followed a systematic rule-and-context protocol and preceded any multi-rater human consensus adjudication. The classifications are therefore interpreted with strict epistemic restraint, and boundary-ambiguous items were flagged for sensitivity testing.[27–30] Knowledge Dimension assignments were also partly coupled to the process rules. Quantitative C3 execution, for example, mapped strongly to Procedural knowledge. The two dimensions are thus not statistically independent cross-validation instruments. Units without defensible context were never forced into C1 Remember. One learning objective in Module 1 used the affective verb 'appreciate' and was kept as undetermined, which leaves 175 codable objectives out of 176. Seven task units with insufficient textual or visual context were excluded from the 494 initial task units, together with the 25 third-party figure-restricted units and five within-type exact duplicates. This leaves 457 unique codable task units: 345 primary comparison tasks [253 homework, 33 quiz, and 59 module examination units], 85 module supplement units, and 27 final examination units.

2.4 Distributional Overlap, Module-Cluster Bootstrap, Sensitivity, and Validation Architecture

The primary alignment analyses compared the 175 codable learning objectives with the 345 primary tasks (253 homework practice units, 33 module quiz units, and 59 module examination units). The 85 supplemental units and 27 final examination units were presented separately and in full-corpus sensitivity models. All six cognitive process categories (C1–C6) and all four knowledge categories were tabulated before any secondary

aggregation into LOTS (C1–C3) and HOTS (C4–C6), and C3 Apply was always displayed apart from C1–C2. Pearson's chi-square (χ^2) tests and Cramér's V effect sizes evaluated two global shifts across the three-tier cognitive partition (C1–C2, C3, C4–C6) and the Knowledge Dimension. One shift ran from learning objectives (n = 175) to primary tasks (n = 345). The other ran from homework practice (n = 253) to combined module assessments (quizzes plus module exams, n = 92).

Module-level taxonomic correspondence was quantified without assuming item-to-item independence. Each instructional stage within each of the nine modules was represented as a normalized probability distribution over the 24 cells of the Cognitive Process × Knowledge Dimension matrix (6 × 4 = 24 cells). Take any two stages within module m, where T denotes learning objectives, L denotes homework practice, and A denotes combined module quizzes and module examinations. Their 24-cell distributional overlap (histogram intersection) is the sum of cell-wise minimum proportions across all 24 cells. It is bounded between 0 (completely disjoint taxonomic profiles) and 1 (identical 24-cell distributions). Across the nine modules, the primary summary metrics weighted each module equally (1/9). They report parametric means ± standard deviations (SD), Shapiro–Wilk normality statistics (W), medians with interquartile ranges (IQR), and Wilcoxon signed-rank tests comparing L–A, T–L, and T–A overlaps.

Sampling and compositional variability across modules was quantified with 5,000 non-parametric cluster bootstrap replications. Each replication resampled the nine modules with replacement and gave every sampled module equal weight. The 345 task sub-items were not resampled as if they were independent observations. The resulting 95% percentile intervals (2.5th to 97.5th percentiles) capture cross-module compositional heterogeneity. The influence of individual modules was evaluated with nine leave-one-module-out (LOMO) jackknife runs. Six deterministic sensitivity specifications tested the robustness of task cognitive tier proportions: (1) primary corpus (n = 345), (2) excluding all ambiguity-flagged units (n = 315), (3) including module supplements and the final examination (n = 457), (4) pre-deduplication within task type (n = 350), (5) collapsing sub-questions into one modal cognitive process per parent problem stem (n = 200), and (6) removing exact normalized text repeats across different task types (n = 345). Finally, controlled topic-tag screening across 24 chemistry sub-topics identified candidate within-module pairs. From these, 36 purposive objective–task pairs were traced qualitatively, and nine representative pairs, one per module, are reported in full. A stratified validation sample of 54 task units (six per module) was structured for independent dual-rater reliability calibration (weighted Cohen's κ for Cognitive Process and nominal κ for Knowledge Dimension).

The quantitative estimators governing stage proportions, 24-cell distributional overlap, module-cluster bootstrap intervals, and leave-one-module-out jackknife stability were formalized algebraically as follows. Let $m \in \{1, ..., 9\}$ index the nine thematic modules, and let $s \in \{T, L, A\}$ denote the three primary module-level instructional stages: codable learning objectives (T, total $n_T = 175$), homework practice units (L, total $n_L = 253$), and combined module assessment units (A, pooling 33 quiz units and 59 module examination units, total $n_A = 92$, with total primary tasks $n_{L+A} = 345$). Within module m and stage s, let $n_{m,s}$ denote the number of codable units, and let $c_{m,s}(i, j)$ denote the count of units assigned to Cognitive Process category $i \in \{1, ..., 6\}$ (corresponding to C1 through C6) and Knowledge Dimension category $j \in \{1, ..., 4\}$ (Factual, Conceptual, Procedural, Metacognitive), such that $\sum_{i=1}^{6} \sum_{j=1}^{4} c_{m,s}(i, j) = n_{m,s}$. The normalized cell proportion in the 24-cell taxonomic matrix is $p_{m,s}(i, j) = c_{m,s}(i, j) / n_{m,s}$. For any pair of stages (s, s') within module m, the 24-cell distributional overlap (histogram intersection) is defined as $O_m(s, s') = \sum_{i=1}^{6} \sum_{j=1}^{4} \min\{p_{m,s}(i, j), p_{m,s'}(i, j)\}$, which satisfies $0 \leq O_m(s, s') \leq 1$. Across all nine modules, the primary equal-weight estimator is $\bar{O}(s, s') = (1/9) \sum_{m=1}^{9} O_m(s, s')$, and the module-level C3 escalation contrast is $\Delta C3_m = 100 \times [p_{m,A}(3, \cdot) - p_{m,L}(3, \cdot)]$, with equal-weight mean $\bar{\Delta}C3 = (1/9) \sum_{m=1}^{9} \Delta C3_m$. Because task units are nested within parent stems (260 parents) and thematic modules (9 modules), the 5,000-draw non-parametric cluster bootstrap resampled M = 9 module indices ($m_1^*(b)$, ..., $m_9^*(b)$) with replacement on each replication $b \in \{1, ..., 5{,}000\}$, computing $\bar{O}^*(b)(s, s') = (1/9) \sum_{k=1}^{9} O_{m(k)^*(b)}(s, s')$ and reporting the 2.5th and 97.5th empirical percentiles. Complementarily, the nine leave-one-module-out (LOMO) jackknife replications omitted each module $r \in \{1, ..., 9\}$ in turn to compute $\bar{O}_{(-r)}(s, s') = (1/8) \sum_{m \neq r} O_m(s, s')$ and $\bar{\Delta}C3_{(-r)} = (1/8) \sum_{m \neq r} \Delta C3_m$.

3. Results

3.1 Curriculum Structure, Sample Delimitation, and Module Unit Distribution

The complete curriculum census of MIT OpenCourseWare 3.091SC encompassed 35 instructional sessions, nine thematic modules, 176 explicit session learning objectives, and 62 standalone written task documents spanning 213 pages (Table 1 and Figure 1A–B). Across the 62 written documents, homework assignments accounted for 33 documents (256 initial segmented units), module quizzes accounted for 10 documents (33 initial units), module examinations accounted for nine documents (69 initial units), supplemental module problem sets accounted for nine documents (96 initial units), and the cumulative final examination accounted for one document (34 initial units), totaling 260 parent problem stems and 494 initial segmented task units evaluated under the two-dimensional Revised Bloom's Taxonomy rubric (Table 1, Table 2, and Figure 1A–B).

Applying the pre-specified eligibility rules to the 176 explicit learning objectives identified 175 codable objectives (99.4%) and one undetermined objective (0.6%, which used the affective verb 'appreciate' and was not forced into a cognitive category; Table 1 and Table 2). Across the 494 initial task units, five units (1.0%) were removed as exact normalized text duplicates within the same task type, 25 units (5.1%) were conservatively withheld due to third-party figure restrictions, and seven units (1.4%) were excluded as undetermined due to insufficient context, leaving 457 unique codable task units (92.5% of the initial task units; Table 1 and Figure 1A–B). Within these 457 codable tasks, the primary analytical comparison corpus comprised 345 task units—specifically 253 homework practice units (73.3%), 33 module quiz units (9.6%), and 59 module examination units (17.1%, yielding 92 combined module assessment units)—while the remaining 112 codable units comprised 85 supplemental problem units and 27 cumulative final examination units (Table 1 and Figure 1A–B). All nine thematic modules contained codable objectives, practice tasks, and module assessments, although unit counts varied substantially across modules: codable objectives ranged from 10 (Electronic Materials, Reactions and Kinetics, and Solid Solutions) to 42 (Atomic Structure; mean 19.4 ± 11.5; median 11.0 [IQR: 10.0–27.0]), homework practice units ranged from 8 (Aqueous Solutions and Solid Solutions) to 50 (Atomic Structure and Crystalline Materials; mean 28.1 ± 16.9; median 23.0 [IQR: 17.0–46.0]), and module assessment units ranged from 5 (Amorphous Materials and Solid Solutions) to 16 (Atomic Structure; mean 10.2 ± 4.3; median 8.0 [IQR: 8.0–14.0]; Figure 1C).

Table 1. Curriculum structure, unit segmentation, and analytical sample delimitation of the MIT OpenCourseWare 3.091SC solid-state chemistry corpus.

Curriculum / Analytical Stage	Unit Count (n)	Proportion of Stage	Substantive Definition and Analytical Boundary Criterion
Instructional syllabus sessions	35 sessions	100.0% of course syllabus	Complete instructional sequence organized across nine thematic solid-state chemistry modules.
Thematic self-assessment modules	9 modules	100.0% of modular structure	Spans Atomic Structure (Module 1) through Solid Solutions (Module 9).
Written problem-set and assessment documents	62 documents (213 pages)	33 homework, 10 quizzes, 9 module exams, 9 supplements, 1 final	Standalone written problem sets and examinations (Sessions 7 and 27 contain no separate homework document).
Explicit session learning objectives	176 objectives	175 (99.4%) codable; 1 (0.6%) undetermined	One objective using the affective verb 'appreciate' is retained as undetermined rather than forced into C1 Remember.
Parent problem stems	260 parent items	100.0% of standalone problems	Complete problem prompts; full parent stem context is preserved for every segmented sub-question.
Initial segmented task units	494 task units	100.0% of initial task units	Lettered sub-questions (a–f) or unsplit parent problems evaluated in light of full parent-problem context.
Within-type exact text duplicates removed	−5 task units	−1.0% of initial task units	Normalized for case, whitespace, and punctuation within the same instructional stage.
Third-party figure-dependent units excluded	−25 task units	−5.1% of initial task units	Withheld where task interpretation depended on third-party copyright-restricted figures.
Insufficient-context units excluded	−7 task units	−1.4% of initial task units	Withheld from taxonomic denominator rather than coerced into low-level C1 Remember.
Unique codable task units (full analytical set)	457 task units	92.5% of initial 494 task units	Comprises 345 primary comparison tasks plus 85 module supplement units and 27 cumulative final exam units.
Primary analytical comparison task corpus	345 task units	253 homework (73.3%), 33 quizzes (9.6%), 59 module exams (17.1%)	Primary task set compared against the 175 codable learning objectives (combined module assessments n = 92).

Note: The 494 initial task units represent sub-questions or unsplit parent prompts nested within 260 parent problems rather than statistically independent items. All taxonomic analyses evaluate the 175 codable learning objectives and 457 unique codable task units.

Table 2. Two-dimensional Revised Bloom's Taxonomy operational classification rubric and boundary decision rules for solid-state chemistry objectives and written tasks.

Taxonomic Dimension	Category Code & Label	Operational Definition in Solid-State Chemistry	Boundary Decision Rule and Anti-Misclassification Criterion
Cognitive Process (C1–C6)	C1 Remember	Retrieve chemical terminology, atomic constants, crystal system names, or definitions.	Prompts beginning with 'Name' or 'Identify' that require stoichiometric calculation or electron configuration rules are not coded C1.
Cognitive Process (C1–C6)	C2 Understand	Explain physical-chemical relationships, bonding mechanisms, or interpret structural representations.	Presence of the surface verb 'explain' or 'describe' alone does not elevate an item to higher-order thinking.
Cognitive Process (C1–C6)	C3 Apply	Execute quantitative equations, thermodynamic/stoichiometric algorithms, or taught representational conventions.	Drawing standard Lewis structures, sketching unit-cell planes, or plotting binary phase diagrams from given data is coded C3, not C6.
Cognitive Process (C1–C6)	C4 Analyze	Diagnose systematic patterns from experimental data or integrate multivariable structure–property relations.	Requires non-routine pattern diagnosis (e.g., determining Bravais lattice centering from X-ray diffraction reflection rules); boundary cases flagged.
Cognitive Process (C1–C6)	C5 Evaluate	Critique a chemical model, material selection, or experimental claim against explicit criteria.	The prompt word 'justify' attached to a routine calculation or single-concept explanation does not qualify as C5 Evaluate.
Cognitive Process (C1–C6)	C6 Create	Design an open-ended material processing sequence or novel experimental solution under multiple constraints.	Constrained plotting or routine structure drawing from fixed invariant parameters remains C3 Apply rather than C6 Create.
Knowledge Dimension	Factual Knowledge	Isolated chemical symbols, terminology, physical constants, and standard nomenclature.	Assigned when tasks require recalling discrete facts without model interpretation or algorithmic execution.
Knowledge Dimension	Conceptual Knowledge	Theories, atomic and band-structure models, bonding classifications, and phase-equilibrium principles.	Captures comprehension of underlying physical-chemical mechanisms and structure–property relationships.
Knowledge Dimension	Procedural Knowledge	Quantitative calculation algorithms, algebraic derivations, and crystallographic/molecular drawing conventions.	Partially coupled with C3 Apply in quantitative chemistry problem solving; not treated as an independent validation test.
Knowledge Dimension	Metacognitive Knowledge	Explicit reflection on problem-solving heuristics, strategy selection, or self-monitoring of cognitive errors.	Assigned only when written prompts explicitly require learners to evaluate or regulate their own cognitive strategies.

Note: Grounded in Anderson and Krathwohl's two-dimensional Revised Bloom's Taxonomy. Each codable unit is assigned one dominant Cognitive Process category and one dominant Knowledge Dimension category based on the complete problem-stem context.

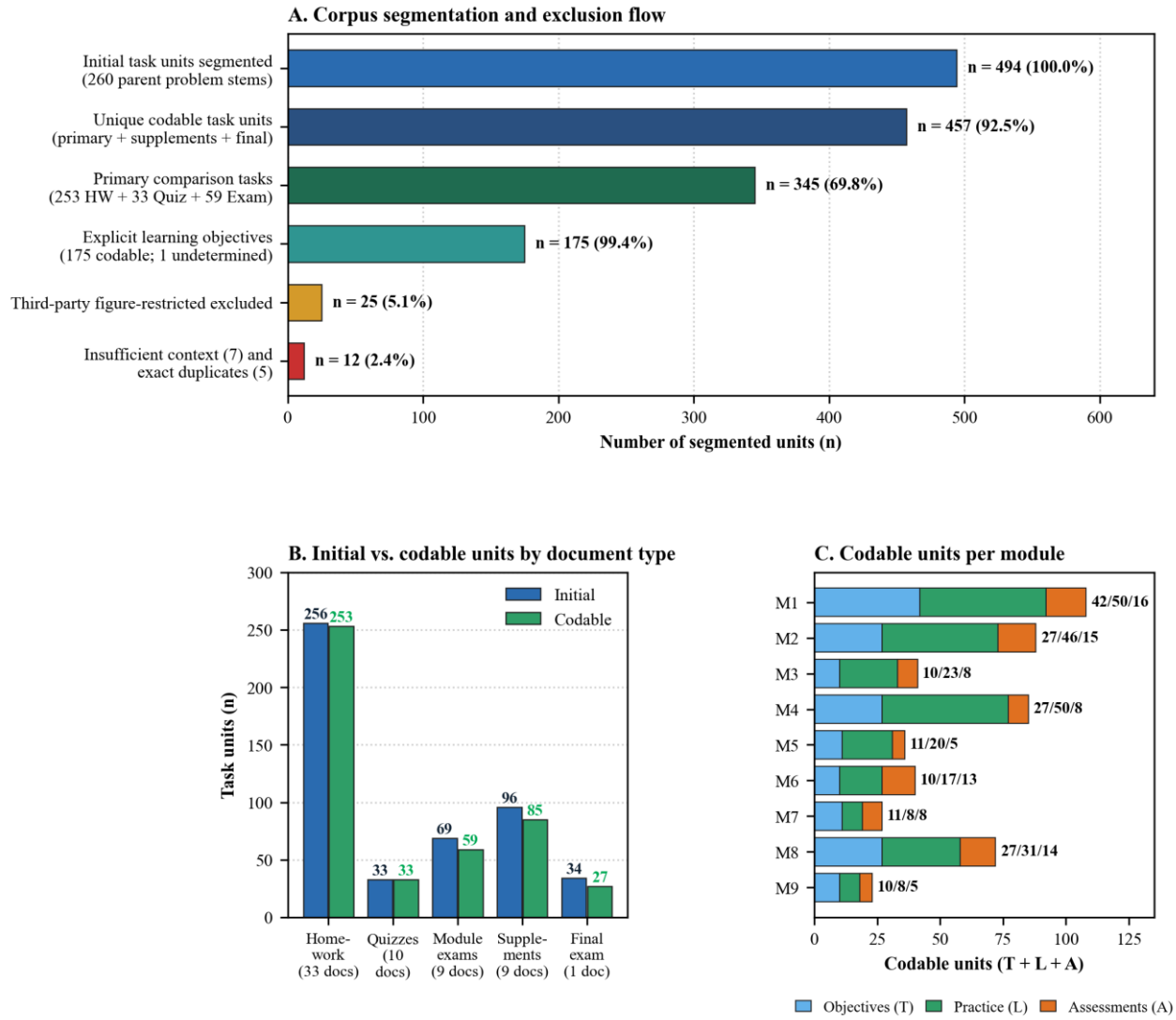


Figure 1. Curriculum structure, unit segmentation, exclusion flow, and nine-module task stratification of the MIT OpenCourseWare 3.091SC Introduction to Solid-State Chemistry corpus: (A) segmentation and eligibility flow from 35 instructional sessions, nine thematic modules, and 62 written task documents (213 pages; 260 parent problem stems; 494 initial labeled task units) through exclusions of five within-type exact duplicates (1.0%), 25 figure-restricted units (5.1%), and seven insufficient-context units (1.4%) to the 457 unique codable task units (92.5%) and the 345 primary comparison tasks (253 homework practice [73.3%], 33 module quizzes [9.6%], and 59 module examinations [17.1%]) paired with 175 codable learning objectives (out of 176 explicit objectives; 99.4%); (B) comparison of initial versus codable task units across the 62 written task documents; and (C) distribution of codable learning objectives (T), homework practice units (L), and module assessments (A) across Modules 1 through 9.

3.2 Cognitive Process and Knowledge Dimension Distributions Across Instructional Stages

Comparing the six Revised Bloom's Taxonomy cognitive process categories across instructional stages revealed a highly pronounced structural divergence between explicit learning objectives and written assessment tasks (Table 3 and Figure 2A–B). Among the 175 codable learning objectives, more than half emphasized C2 Understand (95/175, 54.3%), followed by C3 Apply (50/175, 28.6%), C1 Remember (20/175, 11.4%), C4 Analyze (8/175, 4.6%), C6 Create (2/175, 1.1%), and C5 Evaluate (0/175, 0.0%; Table 3 and Figure 2A). In sharp contrast, across the 345 primary tasks, C3 Apply was the overwhelming modal category at 237/345 units (68.7%), whereas C2 Understand accounted for 88/345 (25.5%), C4 Analyze for 15/345 (4.3%), C1 Remember for 5/345 (1.4%), and both C5 Evaluate and C6 Create for 0/345 (0.0%; Table 3). Disaggregating primary tasks by stage showed that C3 Apply represented 173/253 (68.4%) of homework practice units (with C2 at 65/253 [25.7%], C4 at 11/253 [4.3%], and C1 at 4/253 [1.6%]), rose to 29/33 (87.9%) in module quizzes (with C2 at 3/33 [9.1%] and C4 at 1/33 [3.0%]), and stood at 35/59 (59.3%) in module examinations (with C2 at 20/59 [33.9%], C4 at 3/59 [5.1%], and C1 at 1/59 [1.7%]; Table 3 and Figure 2A). Across the 92 combined module assessments (quizzes plus module exams), C1 accounted for 1/92 (1.1%), C2 for 23/92 (25.0%), C3 for 64/92 (69.6%), and C4 for 4/92 (4.3%). Supplemental problems (n = 85: C2 = 25 [29.4%], C3 = 57 [67.1%], C4 = 2 [2.4%], C5 = 1 [1.2%]) and the cumulative final examination (n = 27: C2 = 8 [29.6%], C3 = 17 [63.0%], C4 = 2 [7.4%]) exhibited closely matching distributions, yielding 311/457 (68.1%) C3 Apply across all 457 codable tasks (Table 3).

Disaggregating Lower-Order Thinking Skills (LOTS) demonstrated why collapsing C1–C3 into a single tier distorts instructional interpretation (Table 3 and Figure 2B). When C1, C2, and C3 were pooled as LOTS, they encompassed 165/175 learning objectives (94.3%) and 330/345 primary tasks (95.7%), creating a superficial appearance of near-identical ~95% LOTS concentration across objectives and tasks. However, separating C1–C2 (remembering and understanding) from C3 (applying) exposed a marked inversion: C1–C2 comprised 115/175 (65.7%) of learning objectives but only 93/345 (27.0%) of primary tasks, whereas C3 Apply jumped from 50/175 (28.6%) in objectives to 237/345 (68.7%) in primary tasks ($\chi^2(2) = 77.92$, $p = 1.20 \times 10^{-17}$, Cramér's V = 0.387; Table 3 and Figure 2B). Conversely, comparing homework practice (n = 253: 27.3% C1–C2, 68.4% C3, 4.3% C4–C6) against combined module assessments (n = 92: 26.1% C1–C2, 69.6% C3, 4.3% C4–C6) revealed virtually identical three-tier proportions ($\chi^2(2) = 0.05$, $p = 0.976$, Cramér's V = 0.012; Table 3).

The Knowledge Dimension mirrored this operational shift from conceptual framing to procedural execution (Table 4 and Figure 3A–B). Learning objectives were dominated by Conceptual knowledge at 101/175 (57.7%), followed by Procedural knowledge at 54/175 (30.9%) and Factual knowledge at 20/175 (11.4%). By contrast, homework practice was dominated by Procedural knowledge at 183/253 (72.3%), with Conceptual knowledge at 66/253 (26.1%) and Factual knowledge at 4/253 (1.6%; Table 4 and Figure 3A). Procedural knowledge reached 30/33 (90.9%) in module quizzes (with 3/33 [9.1%] Conceptual) and 37/59 (62.7%) in module examinations (with 21/59 [35.6%] Conceptual and 1/59 [1.7%] Factual), yielding 250/345 (72.5%) Procedural, 90/345 (26.1%) Conceptual, and 5/345 (1.4%) Factual knowledge across all 345 primary tasks ($\chi^2(2) = 90.05$, $p = 2.79 \times 10^{-20}$, Cramér's V = 0.416 compared against objectives; Table 4 and Figure 3B). Supplemental problems (59/85 [69.4%] Procedural; 26/85 [30.6%] Conceptual) and the final examination (19/27 [70.4%] Procedural; 8/27 [29.6%] Conceptual) maintained the same ~70:30 Procedural-to-Conceptual ratio (328/457 [71.8%] Procedural overall; Table 4). Zero units across all stages were coded as Metacognitive knowledge (0.0%), a finding bounded by written task prompts and dominant-category coding rules.

Table 3. Distribution of the six Revised Bloom's Taxonomy cognitive process categories (C1–C6) and disaggregated Lower-/Higher-Order Thinking tiers across learning objectives and written task stages.

Instructional Stage (Codable Units)	C1 Remember n (%)	C2 Understand n (%)	C3 Apply n (%)	C4 Analyze n (%)	C5 Evaluate n (%)	C6 Create n (%)	LOTS Subtotal: C1–C2 vs. C3 (Total C1–C3 %)	HOTS Subtotal: C4–C6 n (%)
Learning Objectives (n = 175)	20 (11.4%)	95 (54.3%)	50 (28.6%)	8 (4.6%)	0 (0.0%)	2 (1.1%)	65.7% C1–C2 | 28.6% C3 (94.3%)	10 (5.7%)
Homework Practice (n = 253)	4 (1.6%)	65 (25.7%)	173 (68.4%)	11 (4.3%)	0 (0.0%)	0 (0.0%)	27.3% C1–C2 | 68.4% C3 (95.7%)	11 (4.3%)
Module Quizzes (n = 33)	0 (0.0%)	3 (9.1%)	29 (87.9%)	1 (3.0%)	0 (0.0%)	0 (0.0%)	9.1% C1–C2 | 87.9% C3 (97.0%)	1 (3.0%)
Module Exams (n = 59)	1 (1.7%)	20 (33.9%)	35 (59.3%)	3 (5.1%)	0 (0.0%)	0 (0.0%)	35.6% C1–C2 | 59.3% C3 (94.9%)	3 (5.1%)
Combined Module Assessments: Quizzes + Exams (n = 92)	1 (1.1%)	23 (25.0%)	64 (69.6%)	4 (4.3%)	0 (0.0%)	0 (0.0%)	26.1% C1–C2 | 69.6% C3 (95.7%)	4 (4.3%)
Total Primary Tasks: Practice + Assessments (n = 345)	5 (1.4%)	88 (25.5%)	237 (68.7%)	15 (4.3%)	0 (0.0%)	0 (0.0%)	27.0% C1–C2 | 68.7% C3 (95.7%)	15 (4.3%)
Supplemental Module Problems (n = 85)	0 (0.0%)	25 (29.4%)	57 (67.1%)	2 (2.4%)	1 (1.2%)	0 (0.0%)	29.4% C1–C2 | 67.1% C3 (96.5%)	3 (3.5%)
Final Examination (n = 27)	0 (0.0%)	8 (29.6%)	17 (63.0%)	2 (7.4%)	0 (0.0%)	0 (0.0%)	29.6% C1–C2 | 63.0% C3 (92.6%)	2 (7.4%)
All Codable Tasks Combined (n = 457)	5 (1.1%)	121 (26.5%)	311 (68.1%)	19 (4.2%)	1 (0.2%)	0 (0.0%)	27.6% C1–C2 | 68.1% C3 (95.6%)	20 (4.4%)

Note: Percentages are row proportions within each instructional stage. One learning objective using an affective verb ('appreciate') was undetermined, yielding n = 175 codable objectives out of 176. Comparing Learning Objectives (n = 175) against Primary Tasks (n = 345) across the three-tier cognitive partition (C1–C2 vs. C3 vs. C4–C6) yields $\chi^2(2) = 77.92$, $p = 1.20 \times 10^{-17}$, Cramér's V = 0.387; comparing Homework Practice (n = 253) against Combined Module Assessments (n = 92) yields $\chi^2(2) = 0.05$, $p = 0.976$, Cramér's V = 0.012.

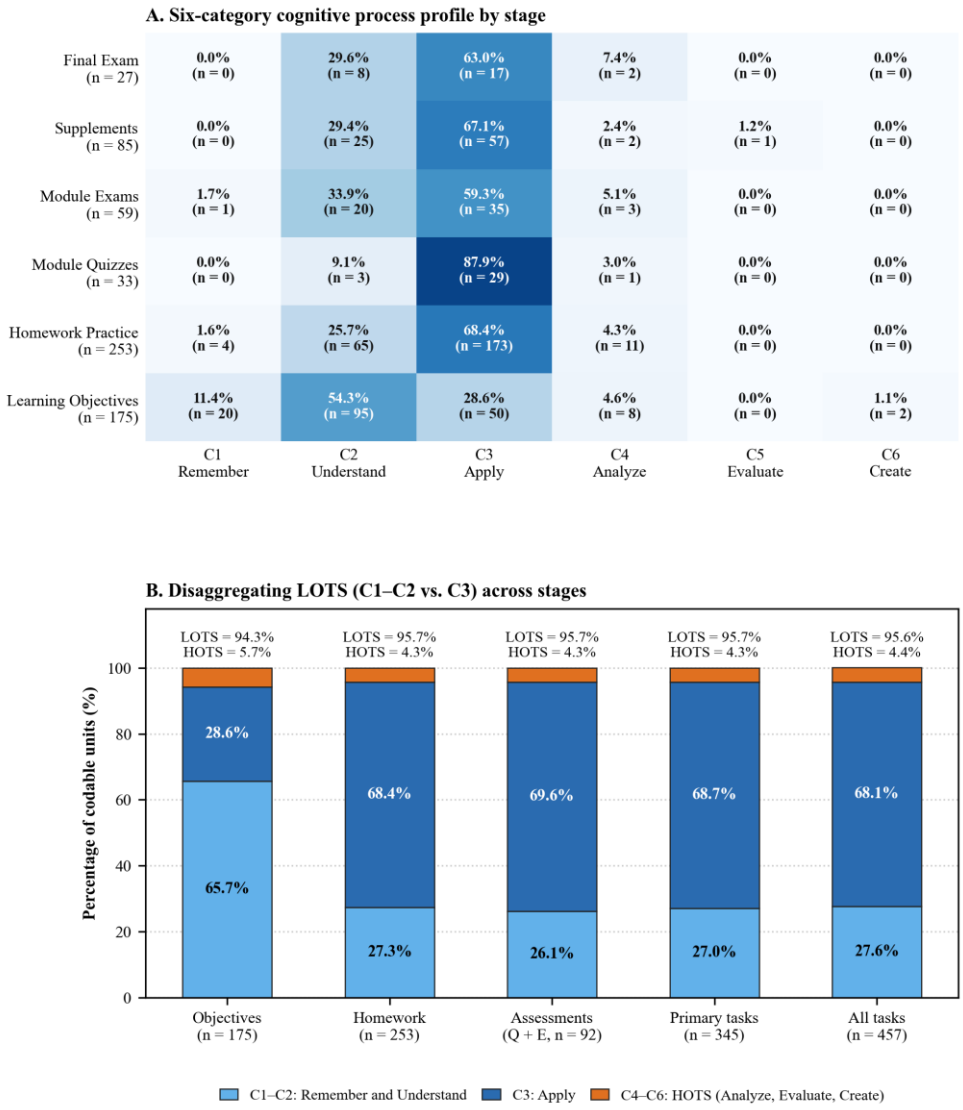


Figure 2. Cognitive process distributions across instructional stages under Revised Bloom's Taxonomy (N = 175 codable learning objectives and N = 457 codable written tasks): (A) complete six-category profile (C1 Remember through C6 Create) across Learning Objectives (n = 175), Homework Practice (n = 253), Module Quizzes (n = 33), Module Examinations (n = 59), Supplemental Problems (n = 85), and the Final Examination (n = 27); and (B) comparison between the coarse binary Lower-Order versus Higher-Order Thinking partition (LOTS C1–C3: 94.3% in objectives vs. 95.7% in primary tasks; HOTS C4–C6: 5.7% vs. 4.3%) and the disaggregated three-tier partition separating introductory comprehension (C1–C2: 65.7% in objectives vs. 27.0% in primary tasks) from procedural application (C3 Apply: 28.6% in objectives vs. 68.7% in primary tasks; $\chi^2(2) = 77.92$, $p = 1.20 \times 10^{-17}$, Cramér's V = 0.387), alongside near-identical distributions between Homework Practice and Combined Assessments ($\chi^2(2) = 0.05$, $p = 0.976$, V = 0.012).

Table 4. Distribution of the four Revised Bloom's Taxonomy Knowledge Dimension categories across learning objectives and written task stages.

Instructional Stage (Codable Units)	Factual Knowledge n (%)	Conceptual Knowledge n (%)	Procedural Knowledge n (%)	Metacognitive Knowledge n (%)	Dominant Knowledge Profile and Pedagogical Interpretation
Learning Objectives (n = 175)	20 (11.4%)	101 (57.7%)	54 (30.9%)	0 (0.0%)	Predominantly Conceptual (57.7%): emphasizes atomic models, bonding principles, and phase relations.
Homework Practice (n = 253)	4 (1.6%)	66 (26.1%)	183 (72.3%)	0 (0.0%)	Predominantly Procedural (72.3%): operationalizes concepts via quantitative calculation and drawing rules.
Module Quizzes (n = 33)	0 (0.0%)	3 (9.1%)	30 (90.9%)	0 (0.0%)	Highly Procedural (90.9%): rapid algorithmic verification of core module calculations.
Module Exams (n = 59)	1 (1.7%)	21 (35.6%)	37 (62.7%)	0 (0.0%)	Procedural majority (62.7%) balanced with Conceptual model explanation items (35.6%).
Combined Module Assessments: Quizzes + Exams (n = 92)	1 (1.1%)	24 (26.1%)	67 (72.8%)	0 (0.0%)	Matches Homework Practice profile (72.8% vs. 72.3% Procedural; 26.1% vs. 26.1% Conceptual).
Total Primary Tasks: Practice + Assessments (n = 345)	5 (1.4%)	90 (26.1%)	250 (72.5%)	0 (0.0%)	Strong shift from Conceptual objectives to Procedural task execution ($\chi^2(2) = 90.05$, $p = 2.79 \times 10^{-20}$, V = 0.416).
Supplemental Module Problems (n = 85)	0 (0.0%)	26 (30.6%)	59 (69.4%)	0 (0.0%)	Consistent Procedural emphasis (69.4%) with Conceptual reinforcement (30.6%).

Instructional Stage (Codable Units)	Factual Knowledge n (%)	Conceptual Knowledge n (%)	Procedural Knowledge n (%)	Metacognitive Knowledge n (%)	Dominant Knowledge Profile and Pedagogical Interpretation
Final Examination (n = 27)	0 (0.0%)	8 (29.6%)	19 (70.4%)	0 (0.0%)	Cumulative synthesis maintaining a ~70:30 Procedural-to-Conceptual ratio.
All Codable Tasks Combined (n = 457)	5 (1.1%)	124 (27.1%)	328 (71.8%)	0 (0.0%)	Zero explicit Metacognitive coding under dominant-category written task rules.

Note: Knowledge Dimension categories were assigned using contextual criteria partially coupled with cognitive process rules; therefore, the two dimensions are not treated as statistically independent cross-validation instruments. Zero Metacognitive counts reflect written task prompts rather than an absence of student self-regulation during study.

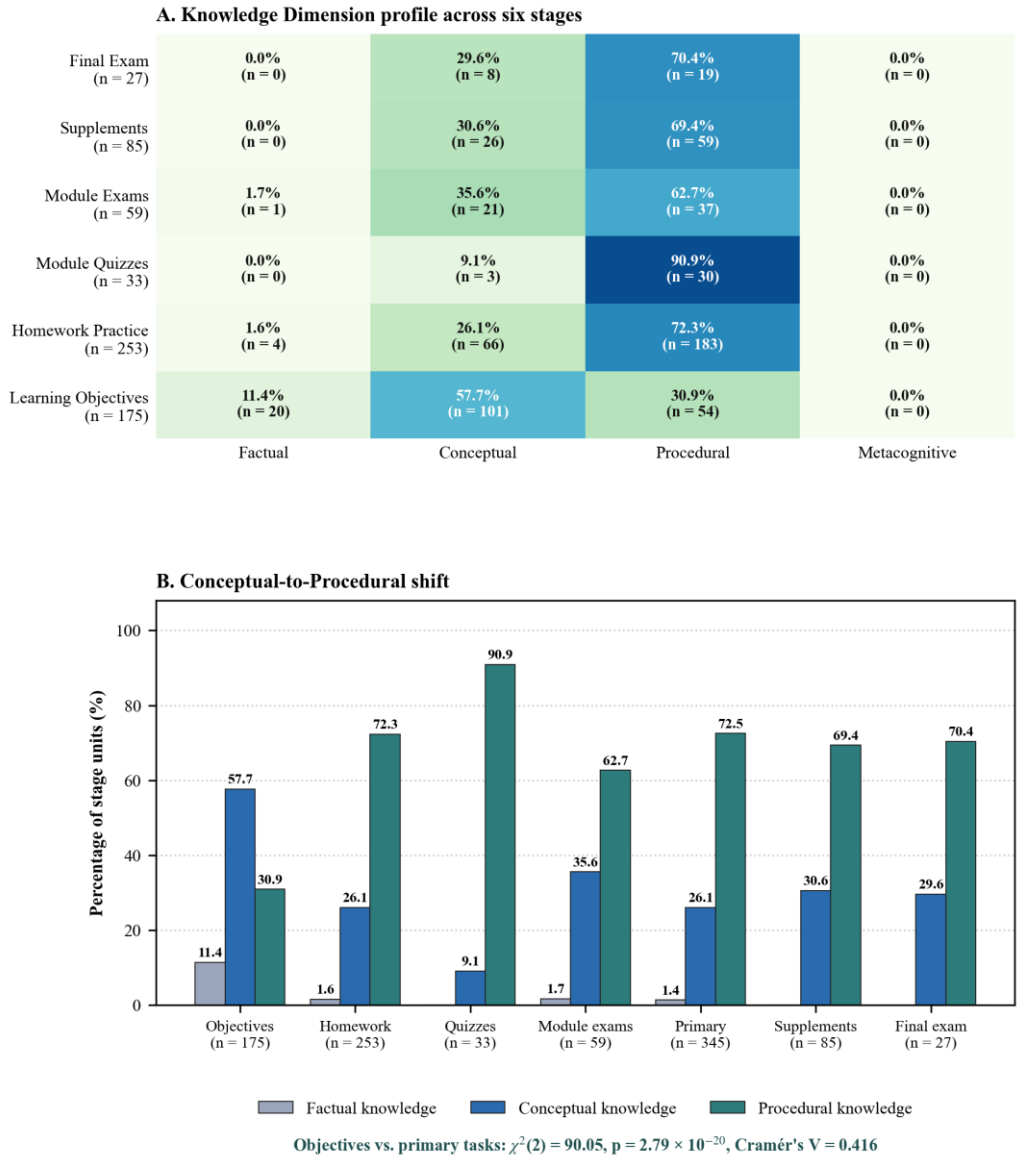


Figure 3. Knowledge Dimension distributions and Conceptual-to-Procedural shift across the MIT OpenCourseWare 3.091SC curriculum: (A) four-category Knowledge Dimension profiles (Factual, Conceptual, Procedural, and Metacognitive Knowledge; 0.0% Metacognitive across all stages) showing the shift from Factual (11.4%) and Conceptual Knowledge (57.7%) in Learning Objectives (n = 175) to Procedural Knowledge in Homework Practice (72.3%), Module Quizzes (90.9%), Module Examinations (62.7%), and Total Primary Tasks (72.5%); and (B) grouped comparison across instructional stages highlighting the Conceptual-to-Procedural operational transition ($\chi^2(2) = 90.05$, $p = 2.79 \times 10^{-20}$, Cramér's V = 0.416).

3.3 Module-Level 24-Cell Distributional Overlap and Resampling Stability

Across the nine thematic modules, 24-cell joint Cognitive Process × Knowledge Dimension distributional overlap varied substantially by chemical topic and instructional pair (Table 5 and Figure 4A–B). Giving equal weight to all nine modules, mean 24-cell overlap between learning objectives and homework practice (T–L) was 0.620 ± 0.206 (median: 0.600 [IQR: 0.527–0.716]; min–max: 0.210–0.922; Shapiro–Wilk W = 0.963, p = 0.832), with a 5,000-resample module-cluster bootstrap 95% percentile confidence interval of 0.490 to 0.743 (Table 5 and Figure 4B). Mean 24-cell overlap between learning objectives and module assessments (T–A, pooling module quizzes and module examinations) was 0.595 ± 0.157 (median: 0.614 [IQR: 0.495–0.696]; min–max: 0.362–0.831; Shapiro–Wilk W = 0.950, p = 0.691), with a 95% bootstrap confidence interval of 0.498 to 0.690 (Table 5 and Figure 4B). In contrast, 24-cell overlap between homework practice and module assessments (L–A) was consistently high across modules, averaging 0.837 ± 0.108 (median: 0.833 [IQR: 0.769–0.922]; min–max: 0.666–1.000; Shapiro–Wilk W = 0.986, p = 0.989; 95% bootstrap CI: 0.769–0.903; Wilcoxon signed-rank test L–A vs. T–A: W = 1.0, p = 0.0078; L–A vs. T–L: W = 7.0, p = 0.0742; T–L vs. T–A: W = 15.0, p = 0.6744; Table 5 and Figure 4B).

Module-by-module inspection identified the specific loci of taxonomic divergence and convergence (Table 5 and Figure

4A). In Module 1 (Atomic Structure; T = 42, L = 50, A = 16), Objective–Practice overlap was the lowest in the course at 0.210 and Objective–Assessment overlap was 0.378 (whereas Practice–Assessment overlap was 0.833): session objectives in Module 1 heavily emphasized historical experiments, quantum-mechanical definitions, and qualitative orbital comprehension (C1 and C2), whereas homework and examination tasks predominantly demanded quantitative photon-energy, de Broglie wavelength, and Bohr/ionization calculations (C3 Procedural). Similarly, in Module 8 (Organic Materials; T = 27, L = 31, A = 14), Objective–Practice overlap was 0.697 but Objective–Assessment overlap dropped to 0.362 (with L–A = 0.666), reflecting a higher concentration of C3 quantitative and structural application items in the module quizzes and examination relative to conceptual objectives, alongside conservative exclusion of third-party figure-dependent organic prompts. Conversely, high Objective–Practice overlap occurred in Module 3 (Electronic Materials: T–L = 0.922, T–A = 0.650, L–A = 0.728), Module 5 (Amorphous Materials: T–L = 0.818, T–A = 0.727, L–A = 0.800), and Module 2 (Bonding and Molecules: T–L = 0.716, T–A = 0.696, L–A = 0.922), while Module 6 (Reactions and Kinetics) achieved the highest Objective–Assessment overlap (T–A = 0.831, T–L = 0.600, L–A = 0.769) and Module 9 (Solid Solutions) exhibited identical Practice and Assessment distributions (L–A = 1.000, T–L = 0.600, T–A = 0.600; Table 5 and Figure 4A). Finally, across the nine equally weighted modules, the mean difference in C3 Apply proportion between module assessments and homework practice (Assessment C3% minus Practice C3%) was +0.6 percentage points, with a 5,000-draw module-cluster bootstrap 95% confidence interval spanning −10.6 to +13.0 percentage points (Table 5 and Figure 4B). Because this interval broadly straddles zero and module-level differences change sign across topics, there is no statistically stable evidence of a course-wide escalation in C3 Apply demand from homework practice to summative module assessments.

Table 5. Module-level distribution of codable units, 24-cell joint Cognitive Process × Knowledge distributional overlap across the nine thematic modules, and module-cluster bootstrap resampling summary.

Module Order & Thematic Title	Total Module Units (n)	Objectives (T, n)	Practice (L, n)	Assessments (A, n)	Supplements (S, n)	Overlap: T vs. L	Overlap: T vs. A	Overlap: L vs. A
Module 1: Atomic Structure	119	42	50	16	11	0.210	0.378	0.833
Module 2: Bonding and Molecules	102	27	46	15	14	0.716	0.696	0.922
Module 3: Electronic Materials	47	10	23	8	6	0.922	0.650	0.728
Module 4: Crystalline Materials	110	27	50	8	25	0.527	0.495	0.940
Module 5: Amorphous Materials	38	11	20	5	2	0.818	0.727	0.800
Module 6: Reactions and Kinetics	56	10	17	13	16	0.600	0.831	0.769
Module 7: Aqueous Solutions	34	11	8	8	7	0.489	0.614	0.875
Module 8: Organic Materials	81	27	31	14	9	0.697	0.362	0.666
Module 9: Solid Solutions	29	10	8	5	6	0.600	0.600	1.000
Cumulative Final Examination	27	—	—	27 (Final)	—	—	—	—
Equal-Weight 9-Module Summary: Mean ± SD	632 codable units	175 total (19.4 ± 11.7)	253 total (28.1 ± 17.0)	92 total (10.2 ± 4.3)	85 codable	0.620 ± 0.206	0.595 ± 0.157	0.837 ± 0.108
Equal-Weight 9-Module Summary: Median [IQR]	—	11.0 [10.0–27.0]	23.0 [17.0–46.0]	8.0 [8.0–14.0]	9.0 [6.0–14.0]	0.600 [0.527–0.716]	0.614 [0.495–0.696]	0.833 [0.769–0.922]
5,000 Module-Cluster Bootstrap 95% Percentile CI	9 module clusters	—	—	—	—	0.490–0.743	0.498–0.690	0.769–0.903

Note: Distributional overlap is computed as $\sum \min(p_c, q_c)$ *across all 24 cells (6 cognitive processes × 4 knowledge types) within each module. Across the 9 equally weighted modules, the module-level C3 Apply difference (Assessment C3% minus Practice C3%) averaged +0.6 percentage points (95% module-cluster bootstrap CI: −10.6 to +13.0 percentage points). [Author query: the S column sums to 96 initial supplemental units, whereas the summary row gives 85 codable supplemental units and 632 codable units; the module rows sum to 616 plus 27 final-examination units; please reconcile].*

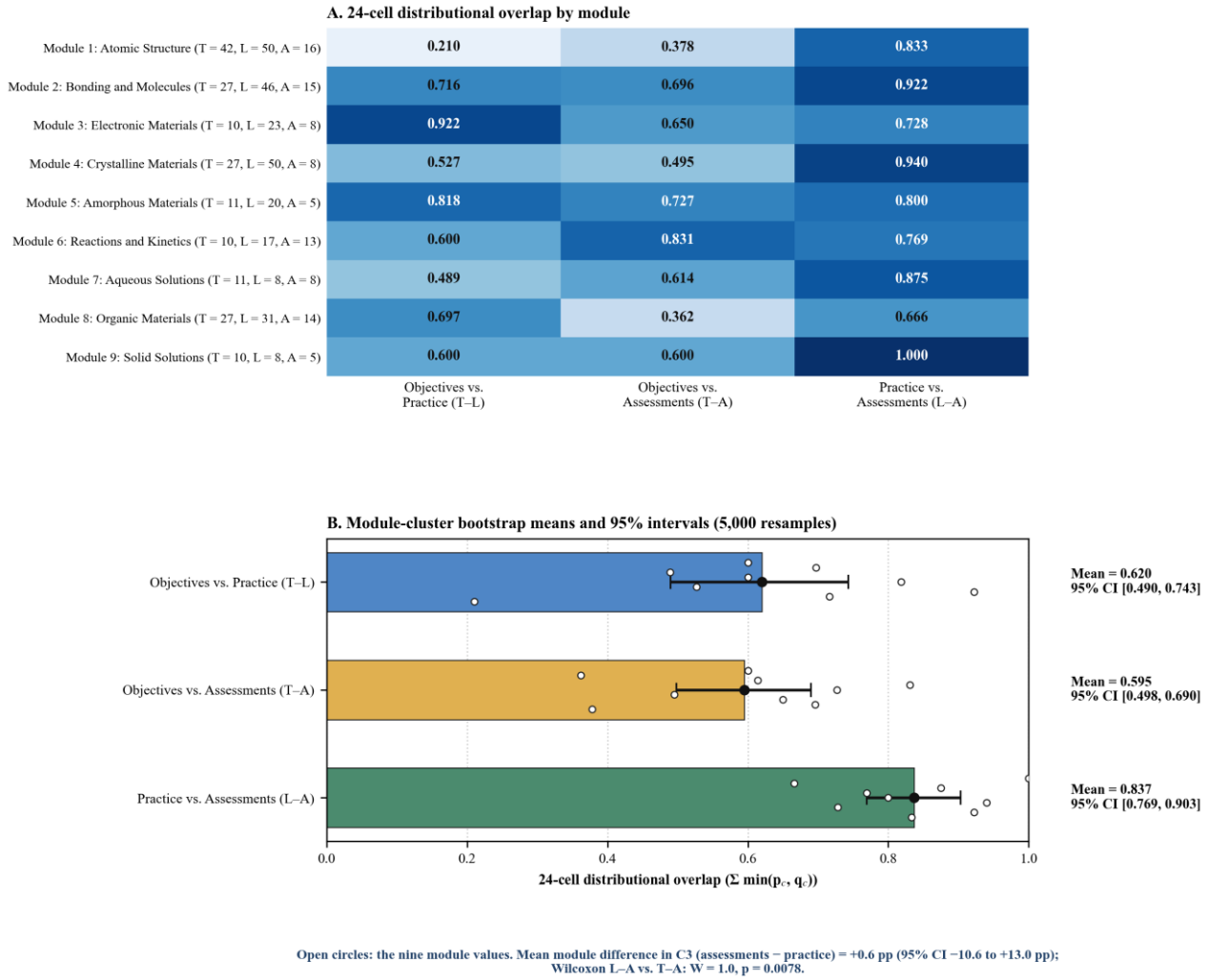


Figure 4. Module-level 24-cell distributional overlap and non-parametric module-cluster bootstrap uncertainty across the nine instructional modules: (A) pairwise 24-cell overlap coefficients for Objectives–Practice (T–L), Objectives–Assessments (T–A), and Practice–Assessments (L–A) across Modules 1 through 9, highlighting high L–A coupling (0.666 to 1.000) and localized divergence in Module 1 (Atomic Structure: T–L = 0.210, T–A = 0.378) and Module 8 (Organic Materials: T–A = 0.362); and (B) equal-weight nine-module mean overlaps with 5,000-replication module-cluster bootstrap 95% percentile confidence intervals: T–L mean = 0.620 ± 0.206 (median 0.600 [IQR: 0.527–0.716]; 95% CI: 0.490–0.743), T–A mean = 0.595 ± 0.157 (median 0.614 [IQR: 0.495–0.696]; 95% CI: 0.498–0.690), and L–A mean = 0.837 ± 0.108 (median 0.833 [IQR: 0.769–0.922]; 95% CI: 0.769–0.903), alongside the module-mean C3 difference between assessments and practice (+0.6 percentage points; 95% CI: −10.6 to +13.0 percentage points).

3.4 Multi-Specification Sensitivity, Jackknife Stability, and Purposive Pair Tracing

Sensitivity testing across six corpus and unit-definition specifications confirmed that C3 Apply dominance is highly robust, whereas Higher-Order Thinking Skills (HOTS, C4–C6) proportions are sensitive to unit granularity and boundary ambiguity (Table 6 and Figure 5A). Relative to the primary analytical corpus (Specification 1, n = 345: 27.0% C1–C2, 68.7% C3, 4.3% C4–C6), excluding all 30 ambiguity-flagged primary units (Specification 2, n = 315) slightly increased C1–C2 to 28.6% and C3 to 69.2% while halving C4–C6 from 4.3% to 2.2%. Expanding the corpus to all 457 codable tasks by adding the 85 module supplement units and 27 final examination units (Specification 3, n = 457) yielded 27.6% C1–C2, 68.1% C3, and 4.4% C4–C6. Retaining the five exact within-type duplicates prior to deduplication (Specification 4, n = 350) produced 27.1% C1–C2, 68.6% C3, and 4.3% C4–C6. Collapsing lettered sub-questions into a single modal cognitive process per parent problem stem (Specification 5, n = 200 parent items) yielded 29.0% C1–C2 and 69.5% C3, but reduced C4–C6 nearly threefold from 4.3% to 1.5%, because localized C4 pattern-diagnosis sub-questions were absorbed into parent problems dominated by C3 calculation sub-items. Specification 6—removing exact normalized text repeats across different task types—left the primary corpus completely unchanged (n = 345: 27.0% C1–C2, 68.7% C3, 4.3% C4–C6), confirming zero verbatim cross-type text duplication while noting that numerical or conceptual variants are not captured by exact string matching (Table 6 and Figure 5A).

Leave-one-module-out (LOMO) jackknife perturbations across the nine modules further demonstrated that global alignment conclusions are not driven by any single outlier module (Table 6, rows 7–8, and Figure 5B). When each of the nine modules was omitted in turn (evaluating the remaining eight modules), mean Objective–Assessment (T–A) 24-cell overlap remained tightly bounded between 0.565 (when omitting Module 6, Reactions and Kinetics) and 0.624 (when omitting Module 8, Organic Materials), around the nine-module baseline of 0.595 (Figure 5B). Similarly, the LOMO mean difference in C3 Apply between module assessments and homework practice ranged from −3.5 percentage points to +3.5 percentage points around the +0.6 percentage point baseline, reversing algebraic sign depending on which module was withheld and reinforcing that global C3 escalation claims are unsupported (Table 6 and Figure 5B).

Purposive qualitative tracing of representative objective–task pairs across all nine modules (selected from 36 candidate pairs identified via 24 controlled chemistry topic tags) illustrated the concrete pedagogical mechanisms linking course objectives to written tasks across the solid-state chemistry curriculum (Table 7 and Figure 6A–B). Direct 1:1 C3 Procedural alignment was evident where quantitative objectives paired with quantitative tasks, including calculating molecular mass

and reaction stoichiometry in Module 1, applying Lewis electron-dot structure conventions in Module 2, calculating extrinsic semiconductor dopant concentrations for target charge-carrier densities in Module 3, applying Fick's diffusion laws to determine thermal processing time in Module 6, and computing aqueous acid–base pH and pOH in Module 7 (Table 7 and Figure 6A). Direct C2 Conceptual alignment occurred in Module 5, where both the learning objective and the homework problem required explaining how alkali glass network modifiers disrupt bridging oxygen connectivity and lower silicate viscosity. Crucially, three pairs highlighted why context-anchored rubric coding outperforms surface verb heuristics: in Module 4, determining crystal structure from X-ray diffraction data required diagnosing systematic Bragg reflection absences (coded C4 Analyze); in Module 8, an objective framed around defining the amino acid isoelectric point (C2 Understand) was operationalized via a quantitative zwitterionic pK_a calculation (C3 Apply); and in Module 9, constructing a binary eutectic phase diagram from fixed invariant data was coded C3 Apply rather than automatic C6 Create (Table 7 and Figure 6A).

Examining the algebraic structure of the six sensitivity specifications (Table 6 and Figure 5A) clarifies why C3 Apply remains invariant (68.1%–69.5%) while Higher-Order Thinking Skills (C4–C6) vary threefold from 1.5% to 4.4%. In Specification 5 (n = 200 parent items), multi-part problems containing lettered sub-questions (a through f) were collapsed by assigning each parent stem the modal cognitive process of its constituent sub-questions. In solid-state chemistry problem sets, a multi-part question on X-ray diffraction or phase equilibria typically opens with two or three quantitative C3 calculation sub-items (such as computing interplanar spacing d_{hkl} or Bragg diffraction angles θ) followed by a single C4 analytical sub-item requiring pattern diagnosis or structural deduction; when aggregated to a single modal label per parent stem, the majority C3 sub-items dominate the parent code, reducing explicit C4–C6 count from 15/345 (4.3%) to 3/200 (1.5%) while C3 Apply remains virtually unchanged at 139/200 (69.5%). Similarly, in Specification 2 (n = 315), excluding the 30 ambiguity-flagged primary units removes items lying near the C2/C3 and C3/C4 interpretive boundaries, reducing HOTS to 7/315 (2.2%) while maintaining C3 at 218/315 (69.2%). This contrast demonstrates that C3 procedural dominance is a structural invariant of the 3.091SC written curriculum, whereas HOTS prevalence depends directly on sub-question segmentation and boundary rules.

Table 6. Sensitivity analysis of task cognitive process distributions across six analytical corpus specifications and leave-one-module-out (LOMO) jackknife stability bounds.

Sensitivity Specification / Jackknife Perturbation	Analyzed Units (n)	C1–C2: Remember & Understand (%)	C3: Apply (%)	C4–C6: Analyze, Evaluate & Create (%)	LOTS Total: C1–C3 (%)	Substantive Stability and Sensitivity Interpretation
1. Primary Analytical Corpus (Baseline)	345	27.0%	68.7%	4.3%	95.7%	Reference specification (253 homework, 33 quizzes, 59 module exams).
2. Excluding All Ambiguity-Flagged Units	315	28.6%	69.2%	2.2%	97.8%	C3 remains dominant (69.2%); HOTS drops from 4.3% to 2.2% when boundary items are removed.
3. Including Module Supplements & Final Exam	457	27.6%	68.1%	4.4%	95.6%	Full codable task corpus shows virtually identical C3 (68.1%) and HOTS (4.4%) proportions.
4. Pre-Deduplication Within Task Type	350	27.1%	68.6%	4.3%	95.7%	Retaining the 5 exact within-type duplicates alters all tier percentages by <=0.1 percentage points.
5. Single Modal Cognitive Process per Parent Stem	200	29.0%	69.5%	1.5%	98.5%	Aggregating lettered sub-items into 200 parent stems preserves C3 (69.5%) but masks sub-item C4 diagnosis (1.5%).
6. Removing Cross-Type Exact Text Repeats	345	27.0%	68.7%	4.3%	95.7%	Zero exact normalized text matches occurred across different primary task types; distribution unchanged.
7. Leave-One-Module-Out (LOMO): T vs. A Overlap Range	9 jackknife runs (8 modules each)	Min: 0.565	Point: 0.595	Max: 0.624	Span: 0.059	Removing any single module shifts mean Objective–Assessment overlap by at most ±0.030.
8. Leave-One-Module-Out (LOMO): Assessment – Practice C3 Δ	9 jackknife runs (8 modules each)	Min: −3.5 pp	Point: +0.6 pp	Max: +3.5 pp	Span: 7.0 pp	Sign reversal across jackknife runs (−3.5 to +3.5 percentage points) confirms absence of global C3 escalation.

Note: pp = percentage points. Across all six unit-definition and boundary-exclusion specifications, C3 Apply remains tightly bounded between 68.1% and 69.5%, whereas HOTS (C4–C6) varies threefold from 1.5% to 4.4%, demonstrating why HOTS point estimates require dual-rater validation.

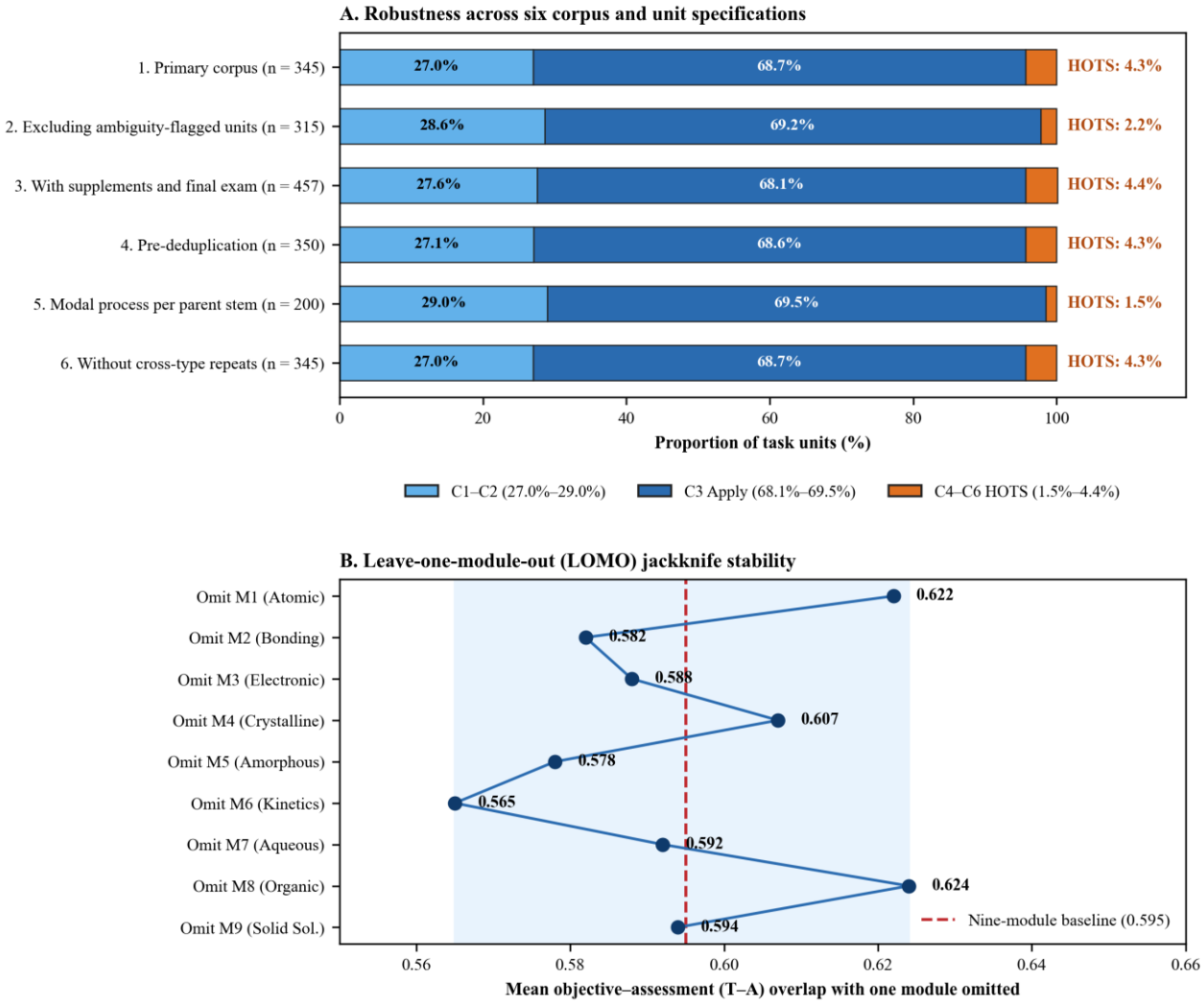


Figure 5. Corpus scope sensitivity and leave-one-module-out (LOMO) stability diagnostics: (A) stability of cognitive process proportions across six analytical specifications—primary tasks (n = 345), excluding ambiguity-flagged units (n = 315), full codable corpus including supplements and final examination (n = 457), pre-deduplication within task type (n = 350), parent-level modal cognitive process per problem stem (n = 200), and removing cross-type exact text repeats (n = 345)—showing invariant C3 Apply dominance (68.1%–69.5%) versus HOTS sensitivity (1.5%–4.4%); and (B) LOMO jackknife trajectory across all nine modules showing that T–A 24-cell overlap remains strictly bounded between 0.565 (omitting Module 6) and 0.624 (omitting Module 8) and that the assessment-minus-practice C3 difference stays bounded between −3.5 percentage points and +3.5 percentage points.

Table 7. Purposive trace of representative objective–task pairs across the nine thematic modules demonstrating direct procedural operationalization, conceptual alignment, and taxonomic shifts.

Thematic Module	Solid-State Chemistry Topic & Task Modality	Objective Taxonomic Code (Process / Knowledge)	Task Taxonomic Code (Process / Knowledge)	Substantive Instructional Alignment Mechanism
Module 1: Atomic Structure	Molecular mass and reaction stoichiometry (Homework practice)	C3 Apply / Procedural	C3 Apply / Procedural	Direct 1:1 procedural alignment: both objective and task require quantitative stoichiometric calculation.
Module 2: Bonding and Molecules	Lewis electron-dot structures and formal charge (Homework practice)	C3 Apply / Procedural	C3 Apply / Procedural	Direct representational alignment: applying taught octet and valence-electron conventions to construct Lewis structures.
Module 3: Electronic Materials	Extrinsic semiconductor doping and carrier density (Homework practice)	C3 Apply / Procedural	C3 Apply / Procedural	Direct quantitative operationalization: computing dopant concentration required to achieve target charge-carrier density.
Module 4: Crystalline Materials	X-ray diffraction indexing and Bravais lattices (Homework practice)	C3 Apply / Conceptual–Procedural	C4 Analyze / Conceptual	Upward analytical operationalization: diagnosing allowed and forbidden Bragg reflection patterns to deduce crystal lattice centering.
Module 5: Amorphous Materials	Alkali network modifiers and silicate glass viscosity (Homework practice)	C2 Understand / Conceptual	C2 Understand / Conceptual	Direct conceptual alignment: explaining how alkali oxide modifiers disrupt bridging oxygen connectivity and lower glass viscosity.
Module 6: Reactions and Kinetics	Solid-state Fickian diffusion and thermal processing (Homework practice)	C3 Apply / Procedural	C3 Apply / Procedural	Direct procedural alignment: applying Fick's diffusion equations and error-function solutions to determine processing time.
Module 7: Aqueous Solutions	Acid–base equilibria, pH, and pOH calculations (Module quiz)	C3 Apply / Procedural	C3 Apply / Procedural	Direct quiz verification: executing logarithmic proton and hydroxide activity calculations in aqueous solution.

Thematic Module	Solid-State Chemistry Topic & Task Modality	Objective Taxonomic Code (Process / Knowledge)	Task Taxonomic Code (Process / Knowledge)	Substantive Instructional Alignment Mechanism
Module 8: Organic Materials	Amino acid zwitterions and isoelectric point pI (Homework practice)	C2 Understand / Conceptual	C3 Apply / Procedural	Constructive operational shift: conceptual definition of zwitterionic charge balance is assessed via quantitative pKa calculation.
Module 9: Solid Solutions	Binary eutectic phase diagram construction (Homework practice)	C3 Apply / Procedural	C3 Apply / Procedural	Rule-governed construction: plotting a binary phase diagram from given invariant temperatures is coded C3 Apply rather than open-ended C6 Create.

Note: These nine representative pairs (one per thematic module) were selected purposively from 36 candidate pairs identified across 24 controlled chemistry topic tags to illustrate qualitative mechanisms of alignment, operationalization, and boundary coding.



Figure 6. Purposive objective-to-task pair alignment trace and four-tier curriculum alignment architecture: (A) cognitive process mapping across nine representative objective–task pairs (one per thematic module), illustrating direct C3 Procedural alignment (Modules 1, 2, 3, 6, 7, and 9), direct C2 Conceptual alignment (Module 5), C2-to-C3 quantitative operationalization (Module 8, isoelectric point calculation), and C3-to-C4 analytical pattern diagnosis (Module 4, X-ray diffraction reflection indexing); and (B) four-tier instructional alignment and validation framework integrating context-preserving segmentation, two-dimensional Revised Bloom's Taxonomy rubric rules, module-cluster resampling, and independent dual-rater reliability calibration (target Cohen's κ ≥ 0.70).

4. Discussion

4.1 Conceptual Objectives versus Procedural Tasks: Operationalization Rather Than Misalignment

The central finding of this analysis is a systematic asymmetry between how learning objectives are written and how written tasks operationalize them across MIT OpenCourseWare 3.091SC. Explicit session learning objectives emphasize C2 Understand (54.3%) and Conceptual knowledge (57.7%), while primary written tasks, which span homework practice, module quizzes, and module examinations, emphasize C3 Apply (68.7%) and Procedural knowledge (72.5%; $\chi^2(2) = 77.92$, $p = 1.20 \times 10^{-17}$ for process; $\chi^2(2) = 90.05$, $p = 2.79 \times 10^{-20}$ for knowledge; Table 3 and Table 4). Within constructive alignment and chemistry education research, this shift calls for a disciplinary reading. It is not automatic evidence of flawed instructional design.[1,5,6,11,12] Instructors in introductory solid-state and general chemistry often phrase session objectives as conceptual understanding, for example comprehending Bohr atomic models, band-gap theory, defect equilibria, or zwitterionic charge states. They then check mastery by asking students to compute photon wavelengths, dopant concentrations, diffusion times, or isoelectric points.[8,9] The Module 8 purposive trace (Table 7) shows the pattern: quantitative C3 application often serves as a legitimate, practical operationalization of C2 conceptual understanding.

The results also show why merging C1, C2, and C3 into one Lower-Order Thinking Skills (LOTS) index misleads in quantitative STEM curricula.[5,6,23] C1–C3 covers 95.7% (330/345) of primary tasks in 3.091SC, so a coarse LOTS/HOTS report could be read as saying that introductory solid-state chemistry relies almost entirely on low-level factual memorization. Disaggregation shows a different picture. C1 Remember accounts for only 1.4% (5/345) of primary tasks, and C1–C2 together for only 27.0% (93/345). A further 68.7% (237/345) demand C3 quantitative or representational application (Figure 2B). In a foundational first-year physical science course for engineering and science undergraduates, extensive procedural practice in crystallography, stoichiometry, thermodynamics, and electronic band calculations is likely to be important. It can build cognitive fluency before students take on open-ended materials design.[22,24]

4.2 Module-Level Distributional Overlap and Absence of Global C3 Escalation

Module-level evaluation uses the 24-cell joint Cognitive Process × Knowledge Dimension distributional overlap. Averages were 0.620 [95% bootstrap CI: 0.490–0.743] for objectives vs. practice, 0.595 [95% CI: 0.498–0.690] for objectives vs. assessments, and 0.837 [95% CI: 0.769–0.903] for practice vs. assessments (Table 5 and Figure 4). This measure guards against misreading distributional indices. A 24-cell overlap of 0.620 or 0.595 describes how similar the category proportions are across the 6 × 4 taxonomic grid within a module, and it does not mean that 62.0% or 59.5% of individual learning objectives were semantically verified as covered by assessment items.[1,23,41] Two stages can show high overlap even when their items address different sub-topics of the same module. A module can also show low overlap, such as Module 1 (Atomic Structure, T–L = 0.210, T–A = 0.378) or Module 8 (Organic Materials, T–A = 0.362). Low values arise when conceptual syllabus definitions are paired with calculation-intensive problem sets, or when third-party figure-dependent organic chemistry questions were conservatively withheld.

For this course, the results do not support a common assumption in curriculum evaluation, namely that summative examinations escalate cognitive process demand toward higher C3 or HOTS proportions relative to formative homework practice.[5,6,12] Across the 345 primary tasks, C3 Apply was highest in short module quizzes (87.9%, 29/33) and lower in module examinations (59.3%, 35/59). C2 Conceptual explanation items made up 33.9% (20/59) of the examination tasks (Table 3). Module quizzes and module examinations were then combined as module assessments (n = 92; 69.6% C3) and compared with homework practice (n = 253; 68.4% C3), giving each of the nine modules equal weight, and the mean difference in C3 proportion was +0.6 percentage points. The 5,000-draw module-cluster bootstrap 95% confidence interval ran from −10.6 to +13.0 percentage points, and the leave-one-module-out range from −3.5 to +3.5 percentage points (Table 5 and Table 6). Homework practice and summative module assessments in 3.091SC thus share a highly consistent taxonomic profile (mean L–A overlap = 0.837; Wilcoxon p = 0.0078 vs. T–A overlap). Students are therefore assessed on a blend of procedural and conceptual demands similar to the one they practice in homework.

The heterogeneity across the nine thematic modules (Table 5 and Figure 4A) shows how subject-matter structure shapes taxonomic overlap in physical science courses.[8,9] Module 1 (Atomic Structure; T = 42, L = 50, A = 16) lists 42 granular learning objectives. They cover the historical discovery of the electron and nucleus, qualitative cathode-ray and gold-foil experiments, and conceptual quantum numbers, which yields high C1 and C2 Conceptual proportions, whereas the six homework problem sets and the module assessments instead concentrate on quantitative photoelectric, Rydberg, de Broglie, and ionization-energy calculations. The result is 0.210 T–L overlap and 0.378 T–A overlap, but a high 0.833 L–A overlap. In Module 8 (Organic Materials; T = 27, L = 31, A = 14), Objective–Practice overlap was 0.697, whereas Objective–Assessment overlap fell to 0.362. The module assessments concentrated on C3 procedural calculations and functional-group transformations, and several visual polymer and biomolecule items with third-party figure restrictions were withheld. Modules built around quantitative physical laws behaved differently. Module 3 (Electronic Materials; T–L = 0.922), Module 5 (Amorphous Materials; T–L = 0.818), and Module 6 (Reactions and Kinetics; T–A = 0.831) showed high 24-cell overlap, because both session objectives and assessment prompts specified quantitative or mechanistic application.

4.3 Actionable Instructional Design Implications for Open Chemistry Courseware

Divergence between C2 objectives and C3 tasks often reflects deliberate operationalization, but the purposive pair tracing and the sensitivity analyses still point to three areas where course designers and OER adopters can strengthen constructive alignment (Table 6, Table 7, and Figure 6B).[1,4,20,21,26,42–45] First, C3 Apply stayed constant across all six sensitivity specifications (68.1%–69.5%). HOTS (C4–C6) was scarce, at 4.3% of primary tasks (15/345 units, all C4 Analyze, with 0.0% C5 and 0.0% C6), and it was fragile. It fell to 2.2% when ambiguous boundary items were removed and to 1.5% when lettered sub-questions were collapsed to one modal process per parent stem (Table 6 and Figure 5A). Syllabus objectives that voiced C6 design aspirations, such as designing a glass processing sequence in Module 5, had no match among primary written tasks coded C6 Create. Phase-diagram construction tasks in Module 9 involved constrained plotting from given invariant data and were coded C3 Apply (Table 7). To cultivate C5 Evaluate or C6 Create in introductory materials chemistry, written assessments would have to go beyond constrained plotting and routine 'justify your calculation' prompts. They would need to require multi-constraint process selection, trade-off evaluation against explicit performance criteria, and open-ended materials synthesis design.[21,26]

Second, none of the 457 codable written tasks received a dominant Metacognitive knowledge code (Table 4). This zero count does not show that students fail to reflect during self-study, but OER self-assessment packages for independent learners could add explicit metacognitive prompts. Learners might be asked to diagnose common unit-conversion errors, compare alternative problem-solving heuristics before computing, or rate their confidence in a chosen conceptual model.[2,3,22] Third, a session-level Objective–Assessment Alignment Matrix could state whether each C2 conceptual objective is assessed through qualitative explanation (C2 Conceptual), quantitative calculation (C3 Procedural), or both. Such a matrix would reduce ambiguity for OER instructors and self-directed students who adapt 3.091SC modules worldwide.[1,13,14]

The following design suggestions are illustrative and were not tested. They indicate how open chemistry courseware could cultivate C5 Evaluate, C6 Create, and Metacognitive knowledge while keeping foundational C3 procedural fluency.[1–4,8,9,21,26,44] In Module 3 (Electronic Materials), a standard C3 Procedural problem asks students to calculate the phosphorus dopant concentration needed for a target electron carrier density in silicon. It can become a C5 Evaluate task if students receive competing dopant candidates with different ionization energies, solid-solubility limits, and diffusivity constants and must then critique the candidates and select the best dopant under explicit thermal-budget criteria. In Module 5 (Amorphous Materials) and Module 9 (Solid Solutions), routine C3 phase-diagram reading can be raised to C6 Create tasks. Students would formulate a multi-stage glass-ceramic annealing schedule or a binary eutectic alloy composition that meets target mechanical toughness, liquidus temperature, and optical transparency constraints at once. Brief metacognitive reflection prompts can sit alongside each problem set. Self-directed learners might state why a

continuum diffusion model applies before substituting numerical values, or identify which limiting-reagent assumption governs their calculation. Prompts of this kind would operationalize Metacognitive knowledge within open self-assessment repositories.[2,3,22]

4.4 Methodological Scope, Limitations, and Dual-Rater Validation Protocol

Four methodological considerations bound the findings of this study (Figure 6B). First, assigning one dominant category can compress multi-step chemistry problems that blend conceptual interpretation with algebraic manipulation. The partial rule-coupling between C3 Apply and Procedural knowledge also means the two dimensions cannot serve as statistically independent validation tests.[27–32] Neither the 5,000-draw module-cluster bootstrap nor the leave-one-module-out jackknife corrects for single-rubric boundary conventions. The three Wilcoxon comparisons on nine modules were also not adjusted for multiplicity.

Second, excluding the 25 third-party figure-dependent units and the seven context-insufficient units was a conservative decision that protects copyright compliance and avoids speculative coding, but it falls more heavily on visually intensive topics such as organic materials and phase diagrams. Third, the corpus is a complete census of one foundational open-access solid-state chemistry course. It leaves out external textbook problems, lecture videos, and classroom grading rubrics, so the findings should not be generalized as a probability sample of global chemistry OER. Fourth, unit counts weight every segmented sub-question equally, with no examination point values, time allocations, or student performance data. Inter-coder reliability across ambiguous boundaries remains to be established. Two independent chemistry education specialists should code the stratified 54-item validation sample (six units per module) and compute quadratic weighted Cohen's κ for Cognitive Process and nominal Cohen's κ for Knowledge Dimension. They should then adjudicate boundary cases between C2/C3 and C3/C4 and calibrate the full-corpus taxonomic estimates.

The planned reliability check uses the stratified 54-unit validation sample (six task units from each of the nine modules) and a four-step dual-rater adjudication workflow.[23,26–32] In Step 1 (Rubric Calibration), two independent domain experts in solid-state chemistry and instructional design review the operational Revised Bloom's Taxonomy rubric (Table 2) with external anchor examples while all initial taxonomic labels are masked. In Step 2 (Independent Dual Coding with Visual Verification), both raters inspect the full problem sets, including embedded crystallographic, phase-equilibrium, and molecular diagrams, with the parent-stem context. Each rater assigns one primary Cognitive Process code (C1–C6), one Knowledge Dimension code (Factual, Conceptual, Procedural, Metacognitive), and a binary secondary-process indicator for multi-step items. In Step 3 (Psychometric Reliability Computation), agreement is quantified before consensus discussion. Quadratic weighted Cohen's κ is used for the six-level ordinal Cognitive Process Dimension, and unweighted nominal Cohen's κ for the four-level Knowledge Dimension, alongside Krippendorff's α and exact cell agreement. The criterion for acceptable reliability is $\kappa \geq 0.70$. In Step 4 (Consensus Adjudication and Corpus Calibration), boundary disagreements are resolved to refine the decision rules and calibrate all corpus distributions. The expected disagreements lie between C2 model explanation and C3 representational convention, and between C3 multi-step calculation and C4 data diagnosis.

5. Conclusion

This two-dimensional Revised Bloom's Taxonomy analysis covered the codable written materials of MIT OpenCourseWare 3.091SC (Introduction to Solid-State Chemistry): 35 instructional sessions, nine thematic modules, 175 codable learning objectives, and 457 codable task units, 345 of them primary tasks. It identifies a pronounced shift. Explicit learning objectives emphasize C2 Understand (54.3%) and Conceptual knowledge (57.7%). Primary written tasks emphasize C3 Apply (68.7%) and Procedural knowledge (72.5%). Separating C3 Apply from C1–C2 (27.0%) within LOTS (95.7%) shows that task demands center on quantitative and representational application, not rote memorization. C3 dominance holds across all six sensitivity specifications (68.1%–69.5%). HOTS (C4–C6: 4.3% baseline; 1.5%–4.4% across sensitivities) depends on sub-question aggregation and boundary filtering. Module-cluster resampling shows no systematic escalation in C3 demand from homework practice to summative assessments (+0.6 percentage points; 95% CI: −10.6 to +13.0). The module-level 24-cell overlaps (0.620 for objectives vs. practice, 0.595 for objectives vs. assessments, 0.837 for practice vs. assessments) and the purposive pair traces give a transparent map for prioritizing curriculum alignment, particularly in Atomic Structure (T–L = 0.210) and Organic Materials (T–A = 0.362). Dual-rater reliability calibration and targeted higher-order task design are the next steps.

Declaration

Artificial Intelligence Use Statement

The authors state that no generative artificial intelligence tool was used to draft or revise the text, to analyze the data, to prepare the figures and tables, or to produce any other scholarly content of this manuscript. The authors take full responsibility for the content of the work.

Author Contribution Statement

Conceptualization, S.A.; methodology, S.A. and K.S.; formal analysis, S.A.; investigation, S.A.; data curation, S.A.; visualization, S.A.; writing – original draft, S.A.; writing – review and editing, S.A., K.S., and N.D.P.; validation, K.S. and N.D.P.; supervision, K.S.; project administration, S.A. All authors read and approved the final manuscript.

Funding Statement

This study was carried out without external financial support. No funding agency in the public, commercial, or not-for-profit sector provided a specific grant for the research, the analysis, or the preparation of this manuscript.

Conflict of Interest / Competing Interests

The authors declare that they have no financial or non-financial relationships that could be regarded as a competing interest in connection with this work. MIT OpenCourseWare 3.091SC was analyzed only as an openly available educational benchmark; the authors have received no commercial endorsement from, and hold no institutional affiliation with, the Massachusetts Institute of Technology.

Ethics Approval and Consent to Participate

Not applicable. This documentary curriculum study examined only publicly accessible open courseware materials (syllabi, session objectives, problem sets, quizzes, and examinations) published by MIT OpenCourseWare under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International license. No human participants, student performance records, classroom observations, or personally identifiable information were collected or analyzed.

Consent for Publication

Not applicable. No identifiable individual participant data, student coursework, or restricted third-party proprietary images are reproduced in this manuscript; 25 task units containing third-party figure restrictions were excluded from the analytical dataset.

Data Availability Statement

All instructional materials analyzed in this study are publicly accessible through the official MIT OpenCourseWare 3.091SC repository (https://ocw.mit.edu/courses/3-091sc-introduction-to-solid-state-chemistry-fall-2010/). The coded dataset, operational coding rubric, 24-cell module contingency matrices, and statistical verification scripts are available from the corresponding author on reasonable request.

Acknowledgments

The authors thank Professor Donald R. Sadoway and the MIT OpenCourseWare team for making the instructional curriculum of Course 3.091SC (Introduction to Solid-State Chemistry) openly available to the global chemistry and materials science education community.

References

1. Biggs J. Enhancing teaching through constructive alignment. High Educ. 1996;32(3):347-364. doi:10.1007/bf00138871
2. Krathwohl DR. A revision of Bloom's taxonomy: an overview. Theory Pract. 2002;41(4):212-218. doi:10.1207/s15430421tip4104_2
3. Crowe A, Dirks C, Wenderoth MP. Biology in Bloom: implementing Bloom's taxonomy to enhance student learning in biology. CBE Life Sci Educ. 2008;7(4):368-381. doi:10.1187/cbe.08-05-0024
4. Towns MH. Developing learning objectives and assessment plans at a variety of institutions: examples and case studies. J Chem Educ. 2010;87(1):91-96. doi:10.1021/ed8000039
5. Parsons AF. Bloom-blueprinted multiple-choice assessment in first-year chemistry: item performance, attainment, and student perceptions. J Chem Educ. 2026;103(4):1612-1621. doi:10.1021/acs.jchemed.6c01082
6. Rylands M, Veale CGL. Adapting Bloom's taxonomy to trace the evolution of assessment patterns in honors organic chemistry. J Chem Educ. 2026;103(5):2569-2574. doi:10.1021/acs.jchemed.5c01371
7. Anderson LC, Fernandez-Branson C. Alignment of learning objectives, assessments, and active learning to promote critical thinking in a first-year medical physiology course: lessons learned. Adv Physiol Educ. 2024;48(2):385-394. doi:10.1152/advan.00096.2023
8. Clark TM, Sample I. Chemistry teachers engagement with solid-state chemistry and materials science concepts through online social annotation. J Chem Educ. 2025;102(12):5169-5177. doi:10.1021/acs.jchemed.5c00575
9. Gai X, Kim MS. Curriculum reform, competency framework, and technological innovation in undergraduate materials science and engineering education: a systematic review. J Chem Educ. 2026;103(4):1793-1803. doi:10.1021/acs.jchemed.5c01454
10. Cracolice MS, Deming JC, Ehlert B. Concept learning versus problem solving: a cognitive difference. J Chem Educ. 2008;85(6):873. doi:10.1021/ed085p873
11. Derman A, Gunes F, Gulacar O, et al. Using a combination technique for the assessment of students' cognitive structures on acid-base chemistry. Chem Educ Res Pract. 2024;25(2):458-473. doi:10.1039/d3rp00142c
12. Shi L, Wang Y, Mitchell-Jones JK, et al. Why do we assess students? Investigating general chemistry instructors' conceptions of assessment purposes and their relationships to assessment practices. Chem Educ Res Pract. 2024;25(4):1159-1174. doi:10.1039/d4rp00147h
13. Sansom RL, Clinton-Lisell V, Fischer L. Let students choose: examining the impact of open educational resources on performance in general chemistry. J Chem Educ. 2021;98(3):745-755. doi:10.1021/acs.jchemed.0c00595
14. Holme T. Open educational resources described in the Journal of Chemical Education. J Chem Educ. 2024;101(3):713-714. doi:10.1021/acs.jchemed.4c00188
15. Niyamae AN, Cheng MMW. Why learning redox is difficult even with animations: students' adherence to cognitive resources. Chem Educ Res Pract. 2026;27(1):662-678. doi:10.1039/d5rp00192g
16. Nakiboglu C, Nakiboglu N. Mapping preservice chemistry teachers' group cognitive structures of electrochemistry and comparison with their understandings of electrochemistry concepts. Chem Educ Res Pract. 2026;27(1):213-240. doi:10.1039/d5rp00128e
17. Zhang L, Wang L, Treagust DF. Discipline-specific cognitive factors that influence grade 9 students' performance in chemistry. Chem Educ Res Pract. 2021;22(4):813-841. doi:10.1039/d0rp00352b
18. Nakiboglu C. Investigation of students' cognitive structures concerning the topic of physical and chemical changes: a cross-level study. Chem Educ Res Pract. 2023;24(1):89-107. doi:10.1039/d2rp00142j
19. Talanquer V. Macro, submicro, and symbolic: the many faces of the chemistry "triplet". Int J Sci Educ. 2011;33(2):179-195. doi:10.1080/09500690903386435
20. Can Aran O, Furtak EM, Buell JY. Supporting chemistry teachers' formative assessment with a three-dimensional learning progression. Int J Sci Educ. 2024;46(15):1601-1622. doi:10.1080/09500693.2023.2291692
21. Abidin ZZ. Learning by brewing tea: student experiences of a Kolb-aligned, constructivist, learning-oriented assessment in chemical engineering. Educ Chem Eng. 2026;54:100495. doi:10.1016/j.ece.2025.100495
22. Braithwaite DW, Sprague L. Conceptual knowledge, procedural knowledge, and metacognition in routine and nonroutine problem solving. Cogn Sci. 2021;45(10):e13048. doi:10.1111/cogs.13048
23. Cerci Akcay H, Ustun N. Cognitive alignment and assessment validity in a high-stakes dental specialty examination: a Bloom's taxonomy-based analysis. BMC Med Educ. 2026;26(1):980. doi:10.1186/s12909-026-09338-w
24. Daher W, Diab H, Rayan A. Artificial intelligence generative tools and conceptual knowledge in problem solving in chemistry. Information. 2023;14(7):409. doi:10.3390/info14070409
25. Livingstone K. A protocol to identify large language model use in undergraduate chemistry essays. J Chem Educ. 2026;103(3):1120-1128. doi:10.1021/acs.jchemed.5c01816
26. Kaldaras L, Yoshida NR, Haudek KC. Rubric development for AI-enabled scoring of three-dimensional constructed-response assessment aligned to NGSS learning progression. Front Educ. 2022;7:983055. doi:10.3389/feduc.2022.983055
27. Anglin KL, Ventura C. Automatic text classification with large language models: a review of openai for zero- and few-shot classification. J Educ Behav Stat. 2025;50(6):1060-1082. doi:10.3102/10769986241279927
28. Kerwat L, Donaldson DL, Mahomed A. Can large language models reproduce higher education grade bands? Cross-model study of calibration and grading bias in authentic student writing. Assess Eval High Educ. 2026;51(4):1-16. doi:10.1080/02602938.2026.2734796
29. Brevik A, Jerpseth H, Lafontan SR. Large language models as grading assistants in public health education: a method-comparison study of essay-style exam assessment. Front Educ. 2026;11:1904452. doi:10.3389/feduc.2026.1904452
30. Yuan Y, Zheng B, Yao C, et al. Multimodal large language model-assisted workflow for image-dependent very short answer question generation and response grading in neurosurgical residency assessment: development and validation study. BMC Med Educ. 2026;26(1):1029. doi:10.1186/s12909-026-10295-7
31. Wei Y, Zhang R, Zhang J, et al. Research on intelligent grading of physics problems based on large language models. Educ Sci. 2025;15(2):116. doi:10.3390/educsci15020116
32. Mok R, Akhtar F, Clare L, et al. Using large language models for grading in education: an applied test for physics. Phys Educ. 2025;60(3):035006. doi:10.1088/1361-6552/adb92b
33. Hailikari T, Virtanen V, Vesalainen M, et al. Student perspectives on how different elements of constructive alignment support active learning. Act Learn High Educ. 2022;23(3):217-231. doi:10.1177/1469787421989160
34. Tobiason G. Going small, going carefully, with a friend: helping faculty adopt lesson-level constructive alignment through non-evaluative peer observation. Act Learn High Educ. 2024;25(1):53-66. doi:10.1177/14697874221092977
35. Salehian Kia F, Pardo A, Dawson S, et al. Exploring the relationship between personalized feedback models, learning design and assessment outcomes. Assess Eval High Educ. 2023;48(6):860-873. doi:10.1080/02602938.2022.2139351
36. Derouich M. Ensuring outcome-based curriculum coherence through systematic CLO-PLO alignment and feedback loops. Discov Educ. 2025;4(1):486. doi:10.1007/s44217-025-00915-7

37. Urrashid H, Ganendra A. Green skills readiness and curriculum alignment in Indonesian vocational schools (SMK): evidence for blue economy and national sustainability competency 2045. Enigma Educ. 2026;3(2):101-111. doi:10.61996/edu.v3i2.115
38. Zorluoğlu SL, Güven Ç. Analysis of 5th grade science learning outcomes and exam questions according to revised Bloom taxonomy. J Educ Issues. 2020;6(1):58. doi:10.5296/jei.v6i1.16197
39. Qadar R, Syam M, Mahdiannur MA. Analyzing high school physics teachers' understanding of cognitive process and knowledge dimensions in assessment design using the revised Bloom's taxonomy. Discov Educ. 2025;4(1):387. doi:10.1007/s44217-025-00807-w
40. Ahmadfard M. Advancing open learning through interactive open educational resources. Discov Educ. 2026;5(1):211. doi:10.1007/s44217-026-01236-z
41. Barthakur A, Joksimovic S, Kovanovic V, et al. Assessing the sequencing of learning objectives in a study program using evidence-based practice. Assess Eval High Educ. 2022;47(8):1429-1443. doi:10.1080/02602938.2022.2064971
42. Cheung KS, Geertshuis S, Liu Q, et al. A systematic curriculum intervention: embedding employability through constructive alignment to enhance student learning. Curric J. 2026;37(2):465-483. doi:10.1002/curj.70016
43. Correia N, Almeida T, Friaes R, et al. Bridging pedagogy, curriculum, and assessment in digital education: ensuring a constructive alignment. Soc Incl. 2025;13:9473. doi:10.17645/si.9473
44. Klapproth F. Redesigning teaching, learning, and assessment in the age of generative AI: a constructive alignment perspective for higher education. Educ Sci. 2026;16(9):1507. doi:10.3390/educsci16091507
45. Calli E, Er E. When technological momentum overshadows pedagogical alignment: a systematic review of AI-generated formative feedback in higher education. Assess Eval High Educ. 2026;51(3):1-24. doi:10.1080/02602938.2026.2674230