International Multidisciplinary Research Journal
ISSN: 2424-7073
Available Online at https://imrjournal.info/
Volume:1; November 2026; Page No.27-38
DOI: https://doi.org/10.47722/imrj.2001.89

Research Article

USING LLMS FOR OUTCOME-ALIGNED COURSE DESIGN

Harry Zhang; Colin Doyle

Embry-Riddle Aeronautical University, U S A

ARTICLE INFO

Article History:

Received 15.07.2026
Accepted 15.10.2026
Published 30.11.2026

Keywords:

outcome alignment, large language models, instructional design, backward design, online learning

ABSTRACT

Designing a high-quality online course requires sustained alignment among course learning outcomes, assessment evidence, module objectives, and learning activities. This paper presents a conceptual framework, illustrated through a worked nine-module course example, for using large language models (LLMs) as alignment-aware design assistants within a backward-design process. The workflow assumes document-grounded use, including retrieval of relevant course-outline and template content, rather than unrestricted generation. It supports three connected stages: clarifying desired learning results, determining acceptable evidence, and planning learning experiences and instruction. Rather than treating AI as a course generator, the paper positions LLMs as tools for interpreting outcomes, proposing aligned design artifacts, checking coherence across stages, and identifying gaps or redundancy across a nine-module course structure. The result is a human-in-the-loop workflow intended to improve traceability, consistency, and efficiency in structured online course development.

Copyright© The Author(s) 2025. Open Access under the Creative Commons Attribution 4.0 International License (CC BY 4.0). This license permits use, sharing, adaptation, distribution, and reproduction in any medium or format, giving proper credit to the original author, including a link to the license, and indicating if any changes were made. To view a copy of the licence, please visit https://creativecommons.org/licenses/by/4.0/

1. Introduction

Designing a high-quality course requires more than assembling content. Effective design depends on alignment among learning outcomes, assessments, module objectives, and learning activities so that students encounter a coherent learning pathway rather than a disconnected set of tasks (Biggs, 1996; Wiggins & McTighe, 2005). In online settings, where learning is mediated through distributed digital materials and activities, this coherence must be made especially visible and maintained across multiple design artifacts (Li et al., 2024). The research gap addressed here is the limited attention to how LLMs can support alignment as a staged design problem rather than as isolated content generation.

Large language models (LLMs) are increasingly discussed as tools for content generation, tutoring, and feedback, but they may also be valuable earlier in the instructional lifecycle. Course design is a text-rich process that requires interpretation, comparison, and revision across outcomes, templates, assessment plans, and module descriptions. These are

areas in which an LLM, when grounded in course documents, may help instructors work more systematically.

This paper argues that LLMs are most useful in course development not as autonomous course builders but as alignment-aware design partners. The paper uses a conceptual synthesis of backward design, constructive alignment, online learning design, and grounded generative AI, then illustrates the framework through a worked example of planning a nine-module online course. The focus is a backward-design workflow in which planning begins with the course outline, proceeds through desired results and acceptable evidence, and culminates in module-level learning experiences across a nine-module course. Within this workflow, the LLM supports drafting, comparison, and consistency checking while human instructors retain pedagogical control.

The paper makes four contributions: it proposes a conceptual framework for LLM-supported outcome-aligned course design; links course outlines, backward-design stages, and module planning in a structured workflow; emphasizes traceability, gap detection, and consistency checking as central AI functions; and illustrates how this approach can support planning across a multi-module online course.

2. Background and Related Work

2.1 Backward Design and Constructive Alignment

Backward design and constructive alignment both argue that course planning should begin with intended learning and then align assessment and instruction to those aims (Biggs, 1996; Wiggins & McTighe, 2005). These perspectives treat learning outcomes as design anchors rather than administrative statements. In multi-module online courses, that principle must be held not only at the course level but also across the sequence of module objectives, activities, and assessments.

2.2 Feedback, Formative Support, and Coherent Learning Design

Feedback research is relevant because it highlights the value of clear expectations, actionable guidance, and coherent learning structures. When outcomes, tasks, and standards are visibly aligned, feedback becomes easier for students to interpret and use (Hattie & Timperley, 2007; Nicol & Macfarlane-Dick, 2006; Carless & Boud, 2018). This concern is particularly important in online learning, where activities, resources, and feedback are often distributed across a learning management system. Recent digital learning-design work likewise treats constructive alignment as a practical method for coordinating objectives, pedagogies, technologies, and learning processes (Li et al., 2024).

2.3 Learning Analytics, Educational Data, and Structured Design Support

Learning analytics has shown the value of structured educational information for decision making, but much of that work focuses on teaching after a course is already running. Comparable needs arise during design, when instructors must compare outcomes, assessment plans, module objectives, and activity descriptions before students ever enter the course (Viberg et al., 2018). This suggests that design artifacts themselves can be treated as analysable educational data.

2.4 Generative AI and LLMs in Higher Education

Research on generative AI in higher education has concentrated on tutoring, content generation, feedback, and instructor productivity. Yet course design is not simply a content-production problem; it is a coherence problem. Recent work applying LLMs to constructive alignment supports the relevance of this direction, while also reinforcing the need to define the instructor's role and the design context explicitly (Pereira et al., 2024). LLMs may be especially valuable when they are grounded in source documents and used to compare artifacts, propose mappings, and surface alignment issues rather than merely generate additional text (Dwivedi et al., 2023; Belkina et al., 2025; Lewis et al., 2020).

2.5 Gap in Literature

Literature therefore leaves an important gap. We have strong theories explaining why alignment matters, growing evidence that AI can support teaching tasks, and technical approaches for grounding model outputs in source materials. What remains underexplored is how LLMs can support a staged, alignment-aware course design workflow that carries course outcomes through assessment planning and module-level instruction.

3. Structured Course Design Context

The framework proposed in this paper is grounded in a structured, module-based approach to course development. It combines backward design and constructive alignment: course learning outcomes establish the intended destination, acceptable evidence defines how achievement will be recognized, and learning experiences prepare students to produce that evidence (Biggs, 1996; Wiggins & McTighe, 2005). This staged logic is useful in online settings because digital resources and activities can otherwise become disconnected from the outcomes they are intended to support.

The course outline or proposal is the formal starting point. It establishes scope, curricular context, and official course learning outcomes. Those outcomes anchor all later design work, but they are often too abstract to specify module objectives, assessment criteria, or activity sequences directly. The outline therefore identifies what matters while the staged design process determines how that intent will be realized.

The workflow then moves through three connected stages. Stage 1 clarifies desired learning results by unpacking what students should understand, know, and be able to do. Stage 2 determines acceptable evidence by specifying assessment forms, performance expectations, and criteria that can credibly demonstrate those results. Stage 3 plans learning experiences and instruction, including activities, materials, practice, feedback, and sequencing that prepare students for the evidence. The sequence is iterative: later review may require revision of an earlier stage.

In a nine-module course, these three stages must also be distributed over time. Each module contains objectives, activities, and assessment-related expectations, but it must contribute to a broader progression toward the course outcomes. Designers therefore need to check coverage, sequencing, cognitive demand, workload, and reinforcement across modules, not only the internal quality of individual activities.

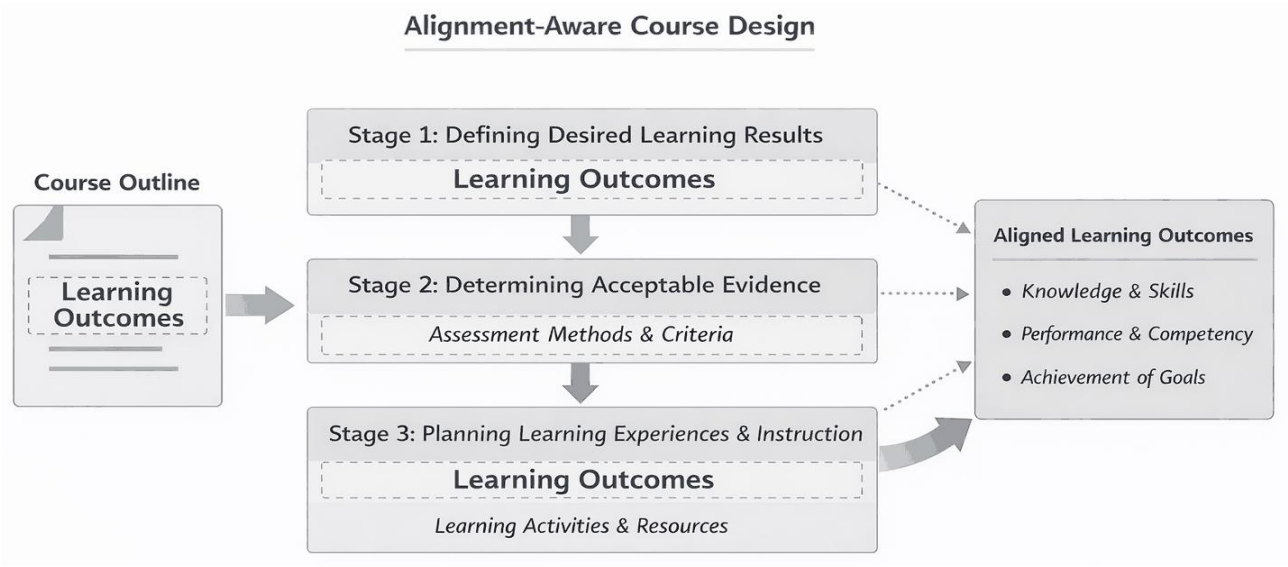


Figure 1 - Alignment of course learning outcomes from the course outline through Stage 1 (Defining Desired Learning Results), Stage 2 (Determining Acceptable Evidence), and Stage 3 (Planning Learning Experiences and Instruction).

This structure is pedagogically significant because a course can appear aligned at the course level yet become uneven when translated into modules. Outcomes may be introduced too late, reinforced too infrequently, or repeated without progression. Module objectives may be locally reasonable but weakly connected to course outcomes, and activities may be selected for convenience rather than their contribution to the expected performance.

These risks explain why LLM support is relevant. Course design involves repeated comparison and revision across outcomes, assessments, objectives, activity instructions, and templates. When grounded in those artifacts, an LLM can help make relationships explicit, identify possible gaps or redundancies, and support review while leaving pedagogical decisions with the instructor.

The framework does not assume that an LLM can determine instructional priorities, disciplinary nuance, student needs, or institutional policy. Its role is narrower: to support the cognitively demanding work of interpreting, mapping, comparing, and revising design artifacts within an established pedagogical process.

The next sections specify the design problems addressed by the framework and show how grounded LLM interactions can support them.

4. Problem Framing: Where LLM Support Is Needed

Interpretation and outcome translation

Broad outcomes must be interpreted before they can guide assessments and module planning. The challenge is to clarify the intended knowledge, skills, and cognitive demand without silently changing the official outcome.

Cross-stage alignment drift

Stage 1 results should shape Stage 2 evidence, and both should inform Stage 3 instruction. In practice, an assessment may capture only part of an outcome, or activities may fail to prepare students for the performance the assessment requires.

Module-level coverage and progression

Across nine modules, designers must check coverage, sequencing, emphasis, and development. Some outcomes may receive too little reinforcement, while objectives may overlap without increasing cognitive demand or independence.

Traceability across artifacts

A module objective should be traceable to a course outcome, acceptable evidence, and the activities that prepare students for that evidence. Without explicit links, quality review and revision become difficult.

Together, these challenges show that useful AI support must extend beyond drafting. A grounded LLM can propose interpretations, generate candidate mappings, compare artifacts, and flag possible gaps or redundancy for human review.

At the same time, not all course design decisions should be delegated to AI. Instructors remain responsible for disciplinary accuracy, instructional priorities, student needs, accessibility, workload, inclusivity, and institutional requirements. The framework therefore treats every model output as a proposal to inspect, revise, accept, or reject.

5. Proposed Framework: LLM-Supported Outcome-Aligned Course Design

This section presents a conceptual framework for using large language models (LLMs) to support outcome-aligned course design within a structured backward-design process. The framework is intended for settings in which course development begins with a formal course outline, proceeds through staged planning artifacts, and culminates in module-level instructional design. Its purpose is not to automate pedagogy or replace instructor expertise. Rather, it positions the LLM as an alignment-aware design assistant that helps instructors and course developers interpret course intentions, draft aligned design artifacts, compare relationships across planning stages, and identify inconsistencies or gaps that may otherwise be difficult to detect. Figure 2 presents the proposed conceptual framework for LLM-supported outcome-aligned course design. The framework begins with pedagogical inputs such as the course outline, learning outcomes, and institutional design templates. These inputs ground LLM support across Stage 1, Stage 2, and Stage 3, while a separate alignment-checking layer helps compare artifacts across stages and modules. Human review remains central throughout the process, ensuring that AI-supported outputs are revised and approved within institutional and pedagogical expectations.

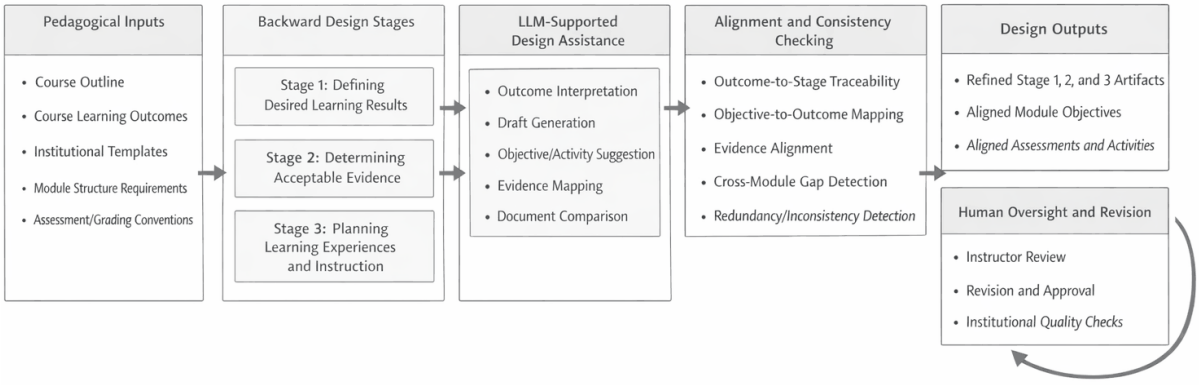


Figure 2 - Conceptual framework for LLM-supported outcome-aligned course design.

At a high level, the framework begins with a set of structured pedagogical inputs, including the course outline, course learning outcomes, stage-based planning documents, and module template requirements. These materials provide the grounding context for all subsequent LLM support. The model is then used to assist with different forms of design work across the planning process: interpreting intended learning, proposing candidate aligned artifacts, checking consistency among those artifacts, and supporting revision through human review. The outputs of the process are not final course materials generated autonomously by AI, but refined design artifacts that remain subject to instructor judgment and institutional expectations.

A key premise of the framework is that meaningful AI support in course planning requires pedagogical grounding. The LLM should not begin with a generic prompt such as "design a course on topic X." Instead, it should operate within a structured instructional context defined by official learning outcomes, design templates, module expectations, and stage-specific planning tasks. This grounding constrains the model's role and makes its outputs more interpretable within institutional course development practices. It also helps shift the use of the LLM away from open-ended content generation toward design support anchored in alignment and traceability.

Example interaction. Given the course outline and Stage 1 results, an instructor could ask: "Using only the attached course outcomes and Stage 1 document, create a traceability table with columns for course outcome, desired learning result, proposed evidence, and supporting module(s). Quote the source outcome, label any inference as a proposal, identify missing evidence, and do not invent requirements. After the table, list the two highest-risk alignment issues for instructor review." The LLM output is useful only as a draft: the instructor verifies every source link, revises proposed evidence, and records accepted changes in the design documents.

In Stage 1, the central task is to clarify what students should understand, know, and be able to do by the end of the course. Although the course learning outcomes provide the official statement of intended learning, they often require further interpretation before they can meaningfully guide design. An LLM can support this stage by helping instructors unpack the outcomes into more explicit formulations of the desired results. This may include identifying major concepts, clarifying the cognitive emphasis of an outcome, proposing enduring understandings, or articulating central questions or themes that give instructional shape to the course.

In Stage 2, the LLM can support the design of acceptable evidence by helping instructors connect learning outcomes and desired results to possible assessment forms, performance tasks, criteria, and indicators of achievement. This may involve proposing ways that students could demonstrate learning, helping distinguish between stronger and weaker evidence of an outcome, or identifying where planned assessments do not fully reflect the intended results established in Stage 1. Retrieval-grounded approaches can further strengthen this process by linking generated suggestions back to course documents and design artifacts (Lewis et al., 2020).

In Stage 3, the LLM can support the design of module objectives, learning activities, instructional sequences, formative practice opportunities, and resource recommendations. In a module-based course structure, this stage is particularly demanding because each module must work both as a local unit of instruction and as part of a larger progression across the full course. In courses that involve technical or programming-oriented artifacts, LLM capabilities for code-related generation and analysis may provide additional support for checking the fit between activities, assignments, and intended learning demands (Jiang et al., 2024).

The most distinctive feature of the framework is its emphasis on alignment checking across the entire design process. While LLMs can certainly assist with drafting at individual stages, their broader value lies in helping instructors compare artifacts that were produced at different points in the process and determine whether those artifacts remain coherent when considered together. This shifts the focus from generation alone to a combination of generation and structured pedagogical analysis.

Although the framework assigns a substantial supportive role to the LLM, it remains firmly human-in-the-loop. Instructor and course designer judgment is essential at every stage, not only because AI outputs may be inaccurate or generic, but because pedagogical design inherently involves contextual, disciplinary, and ethical considerations that cannot be reduced to pattern generation. Decisions about emphasis, appropriateness, inclusivity, workload, sequencing, and instructional tone must remain under human control.

Taken together, the proposed framework conceptualizes LLM-supported course design as a structured process of grounded interpretation, aligned generation, comparative analysis, and human revision. It begins with formal pedagogical artifacts rather than with open-ended prompts. It supports Stage 1, Stage 2, and Stage 3 planning in ways that are sensitive to the instructional purpose of each stage. Most importantly, it uses the LLM not only to create candidate design elements, but also to examine the traceable relationships among learning outcomes, acceptable evidence, module objectives, and learning activities across a full multi-module course.

6. Illustrative Workflow Across a Nine-Module Course

The following worked example shows how the framework can be applied while planning a nine-module online software engineering course. It is illustrative rather than an institutional implementation or a prescribed sequence.

The course outline supplies official outcomes, scope, and constraints. The LLM groups outcome language by learning demand—such as conceptual understanding, analysis, design, or evaluation—without replacing the official wording. Figure 3 provides a visual overview of this workflow.

The instructor then uses the grouped interpretation to clarify Stage 1 desired results, identifies credible Stage 2 evidence, and distributes the results across Stage 3 modules. Early modules may emphasize foundations and guided practice; later modules may emphasize application, synthesis, critique, and integration.

At each transition, the instructor asks whether the proposed artifact preserves the outcome, matches its cognitive demand, and remains feasible for the intended learners. Drafting is followed by comparison and revision rather than by automatic acceptance.

After nine module plans are drafted, the LLM reviews the sequence for missing coverage, redundant objectives, prerequisite problems, abrupt changes in cognitive demand, and weak links between activities and assessments. Table 1 summarizes the inputs, LLM actions, and human review tasks at each stage.

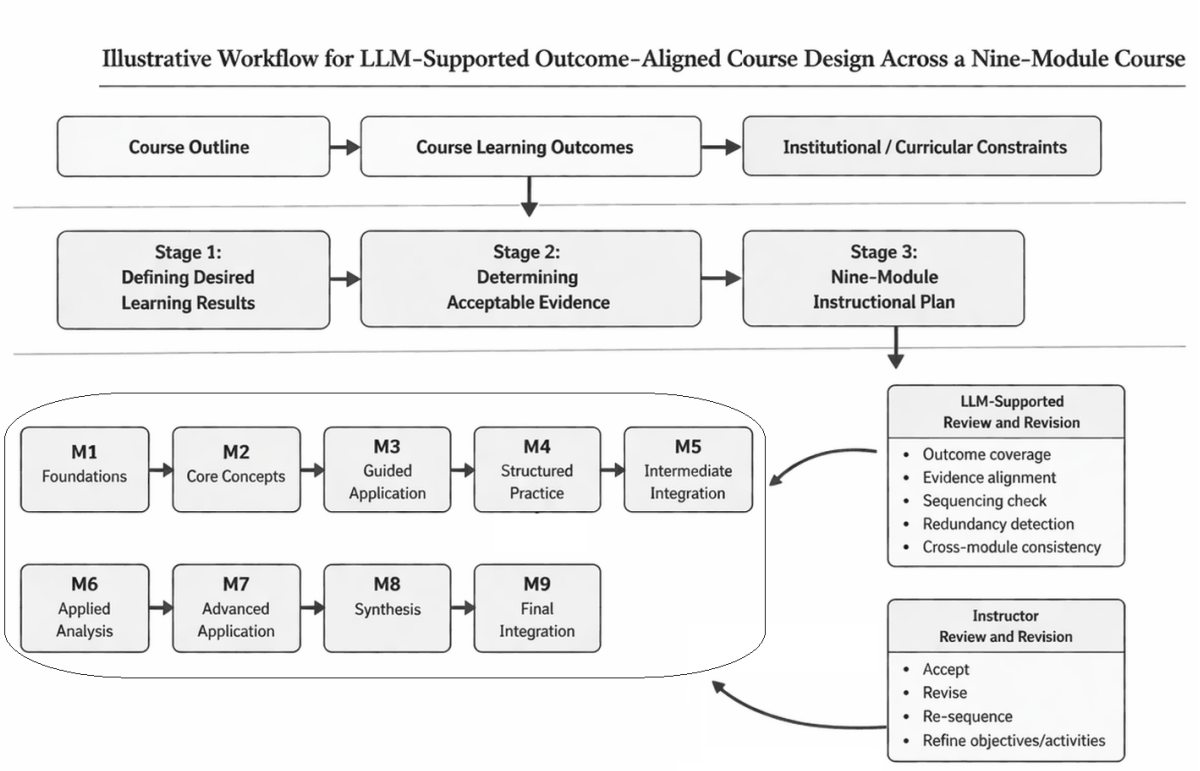


Figure 3 - Conceptual illustration of the Section 6 workflow, in which course-level inputs are translated through Stage 1, Stage 2, and Stage 3 into a nine-module instructional plan that is subsequently refined through LLM-supported and instructor-led review.

Table 1. Summary of the LLM-supported course-design stages

Stage	Primary input	LLM action	Human review task
Course outline	Official outcomes, scope, constraints	Group outcome language and identify learning demands	Confirm interpretation and preserve official wording
Stage 1	Outcomes and design template	Draft desired results and central questions	Check meaning, level, and disciplinary relevance
Stage 2	Desired results and assessment context	Propose evidence, criteria, and performance tasks	Test whether evidence credibly demonstrates each result
Stage 3	Approved results and evidence	Suggest module objectives, activities, practice, and sequence	Check prerequisites, workload, accessibility, and progression
Alignment review	Nine-module plan and traceability links	Flag gaps, redundancy, drift, and weak mappings	Investigate flags and accept, revise, or reject changes

The review is iterative. The instructor investigates each flagged issue, decides whether it is a genuine design problem, and revises the relevant objective, evidence statement, activity,

or cross-reference. The final design is therefore an instructor-approved product supported by AI-assisted analysis.

The example demonstrates that the LLM is most useful when embedded in an existing design process and asked to make alignment relationships visible. Its contribution is not merely faster prose, but structured comparison across stages and modules.

The workflow also makes human oversight operational: instructors verify source grounding, disciplinary appropriateness, cognitive demand, accessibility, workload, and the learner experience before adopting any recommendation.

The result is a repeatable cycle of grounded interpretation, aligned drafting, diagnostic review, and instructor revision across a nine-module course.

7. Design Implications for Practice

Several practical implications follow from this framework. Prompt design should specify the source documents, the design stage, the allowed task, the output schema, and the review criteria. Requests should require the model to quote or identify its source, separate evidence from inference, and state uncertainty. This makes outputs easier to audit than a general request to "improve the course."

Alignment checking should be treated as a central use case. A useful review prompt asks the LLM to compare outcomes, evidence, objectives, and activities against explicit criteria such as coverage, cognitive demand, progression, feasibility, and traceability.

When output is fluent but plausibly misaligned, the instructor should not repair it by intuition alone. The reviewer should identify the specific outcome and evidence involved, check the source documents, label the output as a proposal, and either revise it to restore the missing relationship or reject it. Any accepted change should be recorded in the relevant design artifact.

Novice designers may benefit from scaffolded interpretation and examples, whereas experienced designers may use the system primarily for comparative analysis, gap detection, and revision support. Institutions can use the same traceability record for quality review and later course updates.

8. Benefits, Risks, Limitations, and Future Research

The framework offers several benefits. It can improve coherence across artifacts, make design reasoning more explicit, and reduce some of the time spent drafting and checking alignment manually. It may also help novice designers operationalize backward design while giving experienced designers a useful analytic partner for review and revision.

The risks are equally important. Fluent AI output can mask weak pedagogy, and superficial lexical similarity can be mistaken for meaningful alignment. A model may generate objectives, assessments, and activities that sound related while still failing to reflect appropriate progression or evidentiary strength.

This paper also has clear limitations. It is conceptual rather than empirical, and it is shaped by a structured online course development context that uses formal outlines, stage

documents, and module templates. Different institutional settings may require different workflows, prompts, and review practices.

Future research should evaluate LLM-supported design workflows through mixed-method and comparative studies. A quasi-experimental or matched-course design could compare instructor teams using the grounded workflow with teams using their usual design process. Independent reviewers could score outcome-to-evidence and evidence-to-activity mappings for coverage, cognitive-demand match, progression, and traceability; useful quantitative indicators would include alignment-matrix coverage, weighted gap counts, redundancy rates, inter-rater agreement, revision time, and the proportion of model flags judged valid. Qualitative interviews and think-aloud sessions could examine instructor trust, workload, and how often human reviewers accept, revise, or reject model suggestions. Pilot implementation could then connect design-quality measures to student performance, engagement, and perceptions of coherence while preserving the paper's conceptual scope.

9. Conclusion

The educational value of large language models in course development lies less in rapid text production than in their ability to support alignment work. High-quality course design depends on maintaining coherent relationships among outcomes, evidence, objectives, and learning activities across multiple planning stages and modules.

The framework proposed in this paper responds to that challenge by embedding the LLM within a backward-design workflow grounded in the course outline and other formal design artifacts. Used in this way, the model becomes an alignment-aware assistant that can help instructors interpret outcomes, draft aligned artifacts, and review the coherence of a full multi-module course while preserving human judgment and pedagogical intent. As LLMs develop stronger multimodal and agentic capabilities, future versions of this workflow may be able to inspect diagrams, tables, media, and linked LMS artifacts and coordinate longer review cycles, but those capabilities will increase rather than remove the need for transparent human approval.

10.Artificial Intelligence Tool Disclosure

Artificial Intelligence Tools. The authors used large language model tools to support language editing, organization, and alignment checking during manuscript preparation. The authors reviewed and approved all content and remain responsible for the accuracy, originality, and integrity of the manuscript.

References

Belkina, M., Daniel, S., Nikolic, S., Haque, R., Lyden, S., Neal, P., Grundy, S., & Hassan, G. M. (2025). Implementing generative AI (GenAI) in higher education: A systematic review of case studies. Computers and Education: Artificial Intelligence, 8, Article 100407. https://doi.org/10.1016/j.caeai.2025.100407

Biggs, J. (1996). Enhancing teaching through constructive alignment. Higher Education, 32(3), 347–364. https://doi.org/10.1007/BF00138871

Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. Assessment & Evaluation in Higher Education, 43(8), 1315–1325.

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., ... Wright, R. (2023). So what if ChatGPT wrote it? Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. International Journal of Information Management, 71, Article 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Hattie, J., & Timperley, H. (2007). The power of feedback. Review of Educational Research, 77(1), 81–112. https://doi.org/10.3102/003465430298487

Jiang, J., Wang, F., Shen, J., Kim, S., & Kim, S. (2024). A survey on large language models for code generation. CoRR, abs/2406.00515.

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Kuttler, H., Lewis, M., Yih, W.-t., Rocktaschel, T., Riedel, S., & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. Advances in Neural Information Processing Systems, 33, 9459–9474.

Li, N., Huijser, H., Zhang, S., Zhang, T., et al. (2024). Putting the flesh on the bones: Using a fishbone digital learning design method to align educational objectives. Education and Information Technologies, 29, 25281–25306. https://doi.org/10.1007/s10639-024-12799-5

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. Studies in Higher Education, 31(2), 199–218. https://doi.org/10.1080/03075070600572090

Pereira, E., Nsair, S., Pereira, L. R., et al. (2024). Constructive alignment in a graduate-level project management course: An innovative framework using large language models. International Journal of Educational Technology in Higher Education, 21, Article 25. https://doi.org/10.1186/s41239-024-00457-2

Viberg, O., Hatakka, M., Balter, O., & Mavroudi, A. (2018). The current landscape of learning analytics in higher education. Computers in Human Behavior, 89, 98–110. https://doi.org/10.1016/j.chb.2018.07.027

Wiggins, G. P., & McTighe, J. (2005). Understanding by design (2nd ed.). ASCD.

How to cite this article:
Zhang H.; Doyle (2026) 'Using LLMs for Outcome-Aligned Course Design, *International Multidisciplinary Research Journal*, Volume V; November 2026; Page 27-38.
DOI: https://doi.org/10.47722/imrj.2001.89