Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

The Constructive Alignment Artefact Protocol: A Dual-Instrument Method for Measuring Curriculum Alignment

Craig Sims

craig.sims@curtin.edu.au

Curtin University, School of Education

Method Article

Keywords: constructive alignment, SOLO taxonomy, Porter alignment index, alignment measurement, higher education

Posted Date: March 31st, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9249012/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Abstract

Constructive alignment is widely embedded in higher education curriculum design and quality assurance, yet no standardised quantitative method exists for measuring the degree of alignment across a unit's curriculum components through artefact analysis. This paper introduces the Constructive Alignment Artefact Protocol (CAAP), a dual-instrument coding protocol that integrates the SOLO Taxonomy and Porter's Alignment Index to quantitatively assess constructive alignment through analysis of artefacts before and after a redevelopment intervention. The integration is achieved by adopting SOLO as the cognitive demand dimension of the Porter matrices, creating a shared measurement language across unit learning outcomes, assessment tasks, and teaching and learning activities. The CAAP is demonstrated through a heuristic demonstration using a constructed unit, designed to exhibit known alignment properties in its pre-intervention form and to address these systematically post-intervention. The demonstration shows that the CAAP produces distinct and interpretable outputs across both conditions, detecting non-operational outcome language, cognitive demand gaps, and structural misalignment between curriculum components. The protocol provides a practical, replicable method for quantitative curriculum alignment analysis to support unit design, review, and quality assurance processes in higher education.

Introduction

Constructive alignment, as articulated by Biggs (1996, 2003), holds that effective curriculum design requires coherence between intended learning outcomes, teaching and learning activities, and assessment tasks. Despite its widespread adoption in higher education quality assurance and curriculum review processes internationally (Ashwin et al., 2015), a persistent methodological gap remains: there is no standardised, quantitative method for measuring the degree of constructive alignment in a unit's curriculum design through analysis of its documentary artefacts.

Two established instruments address parts of this problem. The SOLO Taxonomy (Biggs & Collis, 1982) enables classification of the cognitive demand of learning outcome verbs. Porter's Alignment Index (Porter, 2002) provides a quantitative measure of structural coherence between curriculum components. However, these instruments have not previously been integrated, and used separately they produce useful but incommensurable outputs. The integration of SOLO as the cognitive demand dimension within Porter's alignment matrices resolves this limitation, enabling a single unified measurement framework applicable across educational contexts.

This paper introduces the Constructive Alignment Artefact Protocol (CAAP), a dual-instrument coding protocol and accompanying workbook that operationalises this integration. The CAAP enables quantitative measurement of constructive alignment through analysis of curriculum artefacts, specifically unit learning outcome statements (ULOs), assessment task descriptions and rubrics, and teaching and learning activity (TLA) schedules, before and after a curriculum intervention. The protocol is evaluated through a heuristic demonstration using a constructed unit, EDU101 Digital Literacy in

Education, designed to exhibit known alignment properties in its pre-intervention form and to address these systematically in its post-intervention version. The purpose of this demonstration is to establish procedural validity of the protocol rather than to produce empirically generalisable findings (Hevner et al., 2004; Siggelkow, 2007).

This paper addresses the following research question:

How can the SOLO taxonomy be integrated into Porter's Alignment Index to enable a unified quantitative method for measuring constructive alignment through curriculum artefact analysis?

Literature

Constructive Alignment

Constructive alignment, introduced by Biggs (1996) and developed through subsequent work (Biggs, 2003; Biggs & Tang, 2007), is a curriculum design principle holding that effective teaching requires coherence between three components: the intended learning outcomes that specify what students should be able to do, the teaching and learning activities that develop those capabilities, and the assessment tasks that measure whether they have been achieved. The theoretical basis is constructivist: students construct meaning through their own activity, and the assessment task determines what activity students will engage in (Biggs & Tang, 2007). When these components are misaligned, a backwash effect occurs in which students adopt surface learning strategies because the assessment rewards lower-order performance than the outcomes specify (Brabrand & Dahl, 2008).

Constructive alignment is now embedded in higher education quality assurance frameworks, curriculum review processes, and academic development practice internationally (Ashwin et al., 2015). Despite this adoption, the measurement of alignment at the unit level remains largely qualitative and reliant on expert judgement. Existing approaches typically rely on checklist-based audit tools that do not produce quantitative indices of alignment. A standardised, artefact-based method for quantitatively measuring constructive alignment at the unit level has not previously been established.

The SOLO Taxonomy

The Structure of the Observed Learning Outcome (SOLO) Taxonomy, developed by Biggs and Collis (1982), describes five hierarchical levels of learning outcome quality. SOLO 2 and 3, uni-structural and multi-structural respectively, are characterised as quantitative levels in which the student deals with increasing numbers of relevant elements independently. SOLO 4 and 5, relational and extended abstract, are qualitative levels in which the student integrates elements into coherent structures and ultimately generates new perspectives beyond the given information (Brabrand & Dahl, 2008). In the context of constructive alignment, SOLO provides a principled vocabulary for expressing intended learning

outcomes as observable, measurable competences rather than vague aspirational statements (Biggs, 2003).

The application of SOLO to large-scale curriculum analysis was validated by Brabrand and Dahl (2008, 2009) through examination of 632 university course syllabi across Computer Science, Mathematics, and Natural Sciences, with verb classifications confirmed through consultation with Biggs and Tang. Their work established that the operative verbs used in learning outcome statements serve as reliable proxies for the cognitive demand intended by curriculum designers, and that SOLO classification can be applied systematically across disciplines when appropriate calibration procedures are followed. Empirical application in higher education contexts has further demonstrated SOLO's suitability for measuring different kinds of learning outcomes across subjects, levels, and assignment types (Chan et al., 2002).

A critical distinction in SOLO-based curriculum analysis is between operational and non-operational language. Verbs such as "understand," "know," and "be familiar with" describe internal cognitive states that are neither observable nor measurable and cannot be assigned a SOLO level (Brabrand & Dahl, 2008, 2009). Their inclusion would assign numerical weight to outcomes without measurable cognitive demand, undermining the validity of the resulting profile. Operational verbs, by contrast, specify observable performances that can be classified and compared.

Several tensions in the application of SOLO to curriculum analysis warrant acknowledgement. First, verb meanings are not perfectly stable across disciplines: Brabrand and Dahl (2009) note that the same verb can carry different cognitive demands in different disciplinary contexts, requiring calibration to stabilise classifications. Second, treating SOLO levels as equidistant on a numerical scale involves an ordinal-to-interval assumption that is a recognised simplification, though one standard in SOLO-based analyses (Brabrand & Dahl, 2008). Third, Chan et al. (2002) found that conceptual ambiguity at adjacent SOLO levels could be reduced but not eliminated by finer categorisation, a limitation further examined by Stålne et al. (2016), who found that SOLO does not consistently detect subtle differences between outcomes at the same level and that supplementary frameworks may be needed to capture this variation.

These limitations are real but their impact on curriculum alignment analysis is partially mitigated when SOLO is embedded within a structural framework rather than used in isolation. When SOLO serves as the shared cognitive demand dimension across multiple curriculum matrices, the analysis shifts from absolute classification precision to comparative pattern consistency. A unit exhibiting ULOs at SOLO 4 but assessment and teaching concentrated at SOLO 2 reveals a misalignment pattern regardless of whether individual verb classifications are perfectly precise. Classification variability primarily affects the magnitude of observed differences rather than the overall pattern of alignment. The significance of fine-grained level distinctions, the specific limitation identified by Stålne et al. (2016), is also reduced when the analysis operates on distribution patterns across matrices rather than comparisons between individual outcomes. This structural repositioning of SOLO does not eliminate its limitations but meaningfully changes their consequences for alignment measurement.

SOLO analysis in isolation, however, remains unable to assess whether the cognitive demand expressed in learning outcomes is reflected in assessment tasks or teaching activities. That structural question requires a second instrument capable of measuring correspondence across curriculum components.

Porter's Alignment Index

Porter's Alignment Index was developed by Porter (2002) as a quantitative method for measuring the degree of correspondence between curriculum components through content analysis. Originating in standards-based education reform in the United States, the index was designed to assess whether content standards, classroom instruction, and student assessments were conveying consistent messages about what students should learn (Porter, 2002). A key advantage of Porter's methodology is its generalisability: it can be applied to any combination of curriculum components including standards, assessments, instructional materials, and enacted teaching (Fulmer, 2011; Kazemi, 2022). The method has been applied across a range of educational contexts, including studies of teacher instructional practice, curriculum policy, and assessment design (Porter et al., 2007).

The conceptual basis of the index is alignment as structural coherence. Two curriculum components are considered aligned to the degree that they emphasise the same content at the same level of cognitive demand. This is operationalised through a two-dimensional matrix in which content topics form one dimension and levels of cognitive demand form the other. Each component is coded into an identical matrix structure, and the proportional distribution of emphasis across cells is compared. The alignment index summarises this comparison as a single value between 0 and 1.0, where higher values indicate greater correspondence between the two components (Porter, 2002).

$$PAI = 1 - \frac{\sum |X_i - Y_i|}{2}$$

Porter's index has been productively applied across educational levels and contexts, demonstrating sensitivity to meaningful curriculum differences (Porter et al., 2007). Critical values for interpreting PAI scores have been established through Monte Carlo simulation, providing benchmarks against which observed values can be assessed relative to chance (Fulmer, 2011). However, several limitations in the application of the index warrant acknowledgement. Fulmer and Polikoff (2014) demonstrated that PAI alone cannot determine the statistical significance of observed alignment or misalignment, and that supplementary methods are needed to identify whether marginal differences between curriculum components are statistically meaningful. Kazemi (2022) further notes that the PAI value represents only a summary of alignment, and that the underlying proportion matrices are necessary to identify the specific nature and location of any misalignment. These are genuine interpretive constraints that apply to any application of the index.

A more fundamental limitation concerns the cognitive demand dimension. In Porter's original formulation, the cognitive demand categories are defined by researchers according to the subject matter

at hand and are not anchored to a standardised cognitive taxonomy (Kazemi, 2022). In studies of school mathematics and science, these categories typically describe procedural and conceptual levels of performance, but they vary across studies and lack a principled basis for cross-disciplinary comparison. This flexibility, while practical in domain-specific contexts, creates a significant limitation when the method is applied to higher education unit design, where learning outcomes, assessment tasks, and teaching activities span multiple disciplines and where the cognitive demand dimension must be theoretically grounded to be meaningful.

The absence of a standardised, theory-grounded cognitive demand dimension in Porter's original framework is the central limitation that motivates the integration proposed in this paper. What is required is a taxonomy that can consistently classify cognitive demand across curriculum components and disciplinary contexts, and that is theoretically connected to the learning outcome framework already in use in higher education. The SOLO Taxonomy provides a theoretically grounded framework capable of fulfilling this role.

Towards an Integrated Measurement Framework

The preceding review points toward a specific integration opportunity. SOLO and Porter's Alignment Index each address part of the constructive alignment measurement problem, but from different directions and with different limitations. Neither alone can answer the question of whether a unit's intended cognitive demand is structurally supported across its outcomes, assessment, and teaching. No method currently exists that integrates cognitive demand classification and structural alignment measurement into a single artefact-based quantitative framework.

The limitations of each instrument are reciprocal in a theoretically significant way. SOLO produces a cognitive demand profile of intended outcomes but has no mechanism for comparing that profile against assessment tasks or teaching activities. It describes what a unit claims to develop but cannot determine whether that claim is structurally enacted. Porter's index measures structural correspondence across curriculum components but requires a cognitive demand dimension that is not anchored to a standardised theoretical framework (Kazemi, 2022). The two instruments are not merely different tools for the same task; they are addressing different analytical dimensions that are both necessary for a complete measurement of constructive alignment.

The integration proposed in this paper resolves both limitations through a single conceptual move: adopting SOLO as the cognitive demand dimension of the Porter matrices. This aligns the unit of classification used in outcome analysis with the unit of comparison used in structural alignment measurement. The same taxonomy that governs the classification of ULO verbs also governs the cognitive demand axis of the assessment and TLA matrices. SOLO is no longer applied in isolation to a single artefact type; it becomes the shared language through which all three curriculum components are described and compared.

This integration enables four analytical capabilities that neither instrument provides alone: quantitative comparison of cognitive demand across ULOs, assessment tasks, and TLAs within a single framework; detection of misalignment patterns rather than mere description of individual components; consistent pre/post comparison of alignment properties before and after a curriculum intervention; and a unified quantitative representation of constructive alignment through a combination of SOLO distribution metrics and PAI values interpreted together.

The theoretical case for this integration having been established, the following section describes how it is operationalised in practice through the CAAP: the instruments, the protocol, and the heuristic demonstration that tests the framework's diagnostic capabilities against known alignment properties.

Method

Research Design

This study employs document analysis to examine curriculum artefacts as proxies for the intended and planned curriculum (Bowen, 2009). The artefacts examined comprise unit learning outcome statements, assessment task descriptions and rubrics, and teaching and learning activity schedules, representing the three core components of Biggs' constructive alignment framework (Biggs & Tang, 2007). The unit of analysis is the curriculum document rather than the student, and findings speak to the quality of curriculum design rather than to learning outcomes.

A pre/post comparative design is used, in which artefacts from the same unit are examined prior to a curriculum intervention and following redesign, to evaluate changes in alignment properties rather than changes in student outcomes. To demonstrate the protocol's application, this study uses a heuristic demonstration unit, EDU101 Digital Literacy in Education, described in full in the Demonstration Unit section below.

Instrument 1: SOLO Vocabulary Analysis

Instrument 1 classifies the operative verbs in unit learning outcomes using the SOLO Taxonomy, producing a cognitive demand profile of the unit's stated intentions. The classification system follows the verb taxonomy validated by Brabrand and Dahl (2009). Verbs are assigned to SOLO levels 2 through 5, with levels 2 and 3 classified as quantitative competences and levels 4 and 5 as qualitative competences. The full verb reference table is provided in the SOLO Coding Reference Guide available at [https://doi.org/10.17605/OSF.IO/RGTJS].

A critical procedural step is the treatment of non-operational language. Non-operational terms such as 'understand,' 'know,' 'be familiar with,' and 'appreciate' cannot be classified within SOLO and are recorded and excluded from the SOLO calculation (Brabrand & Dahl, 2008, 2009). The proportion of ULOs

containing at least one classifiable verb is reported as a metric indicating the degree to which a unit's stated outcomes are expressed in a form amenable to alignment analysis.

Where a ULO contains multiple operative verbs, each verb is classified independently and the ULO's SOLO score is calculated as the mean of its constituent verb scores. The unit's overall SOLO average is calculated using the double-weight averaging method from Brabrand and Dahl (2008, 2009), in which each ULO contributes equally to the unit average and each verb within a ULO contributes equally to the ULO average. This prevents distortion from ULOs with uneven verb counts.

Three metrics are reported for both pre and post artefacts: the SOLO distribution across levels, the SOLO average, and the proportion of quantitative versus qualitative competences. These metrics provide the cognitive demand dimension used in the alignment matrices of Instrument 2.

Instrument 2: Porter's Alignment Index

Instrument 2 codes curriculum artefacts into two-dimensional matrices and calculates the degree of alignment between pairs of matrices using Porter's Alignment Index. Each matrix has content topics as rows and SOLO taxonomy levels as columns. The use of SOLO as the column dimension directly connects Instrument 2 to the classifications produced in Instrument 1.

Content topics are defined inductively per unit by the coder prior to analysis, derived from the unit artefacts themselves and refined during calibration to ensure categories are consistent, mutually exclusive, and sufficiently specific. Between six and eight topics per unit is recommended to balance sensitivity and sparsity (Porter et al., 2007).

Three artefact types are coded into separate but structurally identical matrices. The intended curriculum is represented by the ULOs, the assessed curriculum by assessment task descriptions and rubrics, and the enacted curriculum by TLA schedules and materials. Each artefact element is assigned a content topic and a SOLO level. Elements are weighted by their relative contribution: ULOs weighted equally, assessment components weighted by mark allocation as a proportion of the total, and TLAs weighted by contact hours as a proportion of total contact time. The cumulative weight in each cell is divided by the total weight across all elements to produce a matrix of proportions summing to 1.0 (Porter, 2002).

The degree of alignment between any two matrices is calculated using Eq. 1,

$$PAI = 1 - \frac{\sum |X_i - Y_i|}{2}$$

where *Xi* and *Yi* represent the cell proportions in the two matrices being compared. The index ranges from 0, indicating no alignment, to 1.0, indicating perfect alignment. Three alignment pairs are calculated

for each unit: ULOs with assessment tasks, ULOs with TLAs, and assessment tasks with TLAs. Each pair is calculated for both pre and post artefacts, producing six PAI values per unit.

To support interpretation, the protocol incorporates critical values established by Fulmer (2011) through Monte Carlo simulation. These values define the range of PAI scores expected by chance for a matrix of a given number of cells. A PAI value above the upper critical value indicates alignment stronger than would be expected by chance. Critical values vary with matrix size and are reported alongside PAI values in the results.

Interpretation of PAI values requires examination of the underlying proportion matrices, which reveal the specific content and cognitive demand combinations where alignment is strong or absent. The PAI value alone does not reveal the nature or location of misalignment.

The Constructive Alignment Artefact Protocol

The CAAP is an integrated coding system that operationalises both instruments within a single unified framework. In design science research, the instantiation of a designed artefact as a working tool demonstrates feasibility and operational implementation (Hevner et al., 2004). The CAAP comprises three interdependent components: a codebook that records all unit-specific coding decisions, a set of procedural calibration steps that establish shared classification conventions before coding begins, and a spreadsheet-based workbook that automates calculation of all outputs from both instruments. Together these components ensure that the theoretical frameworks described in Instruments 1 and 2 are implemented consistently, reproducibly, and with reduced computational burden on the coder or coders, where multiple analysts are used.

The Codebook

Prior to formal coding, the coder familiarises themselves with all available artefacts for the unit to develop an understanding of the content scope before topic categories are finalised. Content topic categories are then defined inductively from the artefacts and recorded in a unit-specific codebook alongside the SOLO verb classification conventions to be applied throughout coding. The codebook serves as the authoritative reference document for all coding decisions made during the study. Any context-dependent classification decisions, ambiguous artefact interpretations, and discipline-specific verb conventions are recorded here and are binding for the remainder of the coding.

C alibration

Before formal coding begins, the coder or coders complete a calibration phase to establish shared classification conventions, particularly for discipline-specific verb usage. This is standard practice in directed content analysis and is essential for cross-disciplinary studies (Brabrand & Dahl, 2009). Where multiple coders are used, each independently classifies ULO verbs in a sample of approximately 10% of artefacts, compares classifications, discusses disagreements, and reaches consensus on discipline-specific verb meanings. Brabrand and Dahl (2009) found that verbs are used differently across faculties

and that consultation with multiple researchers is necessary to stabilise classification, a finding that directly motivates this step. For TLA coding, where operative verbs are typically absent and SOLO level must be inferred from the cognitive work required of students, coders additionally review ambiguous artefacts together during calibration and record agreed conventions. All calibration decisions are recorded in the codebook and are binding for the remainder of the study.

Where multiple coders are used, inter-rater reliability can be reported for three coding decisions: SOLO verb classification, content topic assignment, and matrix cell assignment. For matrix cell assignment, the Pearson correlation between coders' final proportion matrices provides an overall reliability estimate, with Porter (2002) reporting generalisability coefficients of approximately .70 for two analysts and .80 for four as indicative benchmarks. In the present study, a single coder applied the protocol to the heuristic demonstration unit. Transparency is ensured through full documentation of calibration decisions and coding assumptions in the codebook, with limitations noted in the discussion.

The CAAP Workbook

The CAAP Workbook automates the calculations described in Instruments 1 and 2, reducing the computational burden on practitioners and minimising transcription error in proportion matrix construction and PAI calculation. The workbook comprises five functional sheets. The workbook Codebook sheet records unit details, content topic definitions, and calibration decisions. The SOLO Analysis sheet accepts ULO verb classifications and automatically calculates the SOLO distribution, SOLO average, proportion of operational verbs, and proportion of quantitative versus qualitative competences for both pre and post artefacts. Three Porter Matrix sheets accept coded artefact elements for ULOs, assessment tasks, and TLAs respectively, applying the specified weighting rules and generating normalised proportion matrices automatically. The Results Dashboard aggregates outputs from all sheets, displaying pre and post SOLO metrics, PAI values for all three alignment pairs, and Fulmer (2011) critical value benchmarks for the relevant matrix size, enabling direct visual comparison of pre and post alignment profiles. Coders enter artefact elements using structured dropdown menus for content topic and SOLO level assignment, ensuring classification consistency across all matrices.

The full CAAP protocol, SOLO Coding Reference Guide, workbook, and test data guide are openly available to support replication and adaptation at [https://doi.org/10.17605/OSF.IO/RGTJS].

Demonstration Unit

EDU101 Digital Literacy in Education is a constructed first-year undergraduate unit used to demonstrate the CAAP's application and diagnostic outputs. The unit was designed specifically for this purpose, allowing known alignment properties to be embedded in the pre-intervention artefacts and addressed systematically in the post-intervention version.

The pre-intervention version exhibits four properties typical of poorly aligned units: non-operational ULO language, cognitive demand mismatches between ULOs and assessment tasks, concentration of teaching activity at lower SOLO levels, and gaps in content coverage. The post-intervention version

redesigns ULOs, assessment tasks, and TLAs to produce stronger alignment across all three artefact types.

The demonstration shows that the CAAP can detect and quantify these differences. It does not claim generalisability or empirical validation. Those claims require subsequent empirical application, which the openly available protocol is designed to support. The full EDU101 artefacts, codebook, and test data guide are available at [https://doi.org/10.17605/OSF.IO/RGTJS].

Results

The CAAP was applied to EDU101 Digital Literacy in Education using the pre and post curriculum artefacts shown in Table 1. The artefacts comprise five pre-intervention ULOs and four post-intervention ULOs, a single final examination pre-intervention replaced by two assessed tasks post-intervention, and ten one-hour lectures pre-intervention replaced by a mixed schedule of five lectures and five two-hour workshops post-intervention. Full artefacts, coded workbook, and codebook are available at [https://doi.org/10.17605/OSF.IO/RGTJS].

Table 1

EDU101 Curriculum Artefacts and Topic Mapping, Pre and Post

Component	Pre-intervention	Post-intervention	Topic(s)
ULOs	Understand digital literacy frameworks for Australian education	Describe and compare digital literacy frameworks relevant to Australian education	Digital literacy frameworks
	Identify key digital tools for classroom use	Apply digital tools to design a learning activity for a specific year level	Digital tools for learning
	Describe online safety considerations for school settings	Analyse online safety risks in educational digital environments	Online safety and ethics
	Evaluate digital content creation approaches for diverse learners	Design and evaluate digital content for diverse learners, applying accessibility principles	Digital content creation; Accessibility and inclusion
	Be familiar with accessibility standards in education		Accessibility and inclusion
Assessment	Final exam (100%): MCQ on frameworks and tools; short answer on safety and content creation	Task 1 (40%): Portfolio, design activity with tools and analyse safety; Task 2 (60%): Report, compare frameworks, design content, evaluate accessibility	All topics
TLA	10 × 1-hour lectures: frameworks, tools, safety, content, accessibility, revision	5 × 1-hour lectures plus 5 × 2-hour workshops: frameworks, tools, safety, content, accessibility, integration	All topics

These artefacts form the basis for the SOLO classification and Porter alignment analyses reported in the following sections.

Content Topic Codebook

Five content topics were defined inductively from the EDU101 artefacts, with one residual emergent category, producing a matrix of six content topics (including an emergent category) across four SOLO levels (6 × 4 = 24 cells). Fulmer (2011) critical values for a 24-cell matrix at 30 coding points (total weighted entries across the matrix) are 0.787 (lower, $\alpha = .025$) and 0.885 (upper, $\alpha = .975$). These benchmarks are applied to all PAI values reported below.

Table 2

EDU101 Content Topic Codebook

#	Topic	Description
1	Digital literacy frameworks	Conceptual models and policy frameworks for digital literacy
2	Digital tools for learning	Software, platforms, and applications used in educational contexts
3	Online safety and ethics	Student safety, privacy, and ethical use of digital content
4	Digital content creation	Designing and producing digital learning materials
5	Accessibility and inclusion	Universal design, assistive technology, and inclusive digital practice
6	Emergent	Content not captured by topics 1–5

Instrument 1: SOLO Vocabulary Analysis

Table 3 presents the ULO verb classifications for both conditions. The pre-intervention unit contains five ULOs, two of which use non-operational language and are excluded from the SOLO calculation, leaving three operational ULOs with verbs at SOLO levels 2, 3, and 5 and no representation at SOLO 4. The post-intervention unit contains four ULOs, all operational, with seven verbs distributed across SOLO levels 3 through 5, including four verbs at SOLO 4. Two post-intervention ULOs contain multiple verbs and are averaged using the double-weight method prior to unit-level aggregation.

Table 3

ULO Verb Classifications, Pre and Post

ULO	Verb	SOLO Level	Notes
PRE			
1	understand	Non-operational	Excluded
2	identify	2	
3	describe	3	
4	evaluate	5	
5	be familiar with	Non-operational	Excluded
POST			
1	describe	3	Two verbs: ULO average = 3.5
1	compare	4	
2	apply	3	Two verbs: ULO average = 3.5
2	design	4	
3	analyse	4	
4	design	4	Two verbs: ULO average = 4.5
4	evaluate	5	

The SOLO metrics are reported in Table 4. The SOLO average increased from 3.33 to 3.86, a shift of + 0.53. The proportion of operational ULOs increased from 0.60 to 1.00 and the proportion of qualitative competences increased from 33.3% to 71.4%. The most structurally significant change in the ULO set is the shift from no SOLO 4 representation pre-intervention to four SOLO 4 verbs post-intervention, filling the gap in mid-level relational demand that was absent from the pre-intervention profile.

Table 4

SOLO Metrics, Pre and Post

Metric	Pre	Post	Change
Total ULOs	5	4	-1
Operational ULOs	3 (60%)	4 (100%)	+ 40%
SOLO 2 verbs	1 (33.3%)	0 (0%)	-1
SOLO 3 verbs	1 (33.3%)	2 (28.6%)	-1
SOLO 4 verbs	0 (0%)	4 (57.1%)	+ 4
SOLO 5 verbs	1 (33.3%)	1 (14.3%)	0
SOLO average	3.33	3.86	+ 0.53
Quantitative (S2–S3)	66.7%	28.6%	−38.1%
Qualitative (S4–S5)	33.3%	71.4%	+ 38.1%

The absence of SOLO 4 verbs in the pre-intervention ULO set is reflected in the ULO proportion matrix, where no weight appears at SOLO 4 and content coverage is limited to three topics. In the post-intervention matrices, weight is distributed across SOLO 3–5 with representation across all topics.

Instrument 2: Porter's Alignment Index

Table 5 presents the pre-intervention ULO proportion matrix as a worked example of matrix construction. Each of the three operational ULOs contributes equal weight of 0.33, with weight appearing in only three of the 24 cells: Digital tools for learning × SOLO 2, Online safety and ethics × SOLO 3, and Digital content creation × SOLO 5. Digital literacy frameworks and Accessibility and inclusion carry no ULO weight and the SOLO 4 column is entirely empty, directly reflecting the non-operational exclusions and absent SOLO 4 verbs reported in Instrument 1.

Table 5

PRE ULO Proportion Matrix (Instrument 2 Worked Example)

Topic	SOLO 2	SOLO 3	SOLO 4	SOLO 5	Row Total
Digital literacy frameworks	0	0	0	0	0
Digital tools for learning	0.33	0	0	0	0.33
Online safety and ethics	0	0.33	0	0	0.33
Digital content creation	0	0	0	0.33	0.33
Accessibility and inclusion	0	0	0	0	0
Emergent	0	0	0	0	0
Column total	**0.33**	**0.33**	**0**	**0.33**	**1.00**

Note. Cell values are proportions summing to 1.00. Each operational ULO contributes equal weight (1/3). Non-operational ULOs excluded.

Pre-intervention, the assessment matrix concentrates entirely at SOLO 2–3 with no representation above SOLO 3, and the TLA matrix shows the same pattern across all ten sessions. Post-intervention, the ULO matrix distributes weight across SOLO 3–5 with all five topics represented, the assessment matrix distributes across all topics at SOLO 4–5, and the TLA matrix includes both SOLO 3 lectures and SOLO 4–5 workshops across all topics. The shift from a uniform low-level structure pre-intervention to a differentiated higher-order structure post-intervention is consistent across all three artefact types.

Table 6 presents the six PAI values alongside the Fulmer (2011) critical value benchmarks. Pre-intervention, ULO ↔ Assessment (0.50) and ULO ↔ TLA (0.40) both fall below the lower critical value of 0.787, indicating alignment weaker than expected by chance. Pre-intervention Assessment ↔ TLA (0.80) falls just above the lower critical value, reflecting structural consistency between two artefact types both concentrated at SOLO 2–3. Post-intervention, ULO ↔ Assessment increased to 0.70 and ULO ↔ TLA increased to 0.83, with the ULO ↔ TLA value exceeding the lower critical value. Assessment ↔ TLA decreased to 0.73, reflecting divergence between the broader SOLO distribution in TLAs and the more concentrated SOLO 4–5 assessment profile.

Table 6

PAI Values, Pre and Post, with Fulmer (2011) Critical Value Benchmarks

Alignment Pair	Pre PAI	Post PAI	Change	Fulmer Lower	Fulmer Upper
ULO↔ Assessment	0.50	0.70	+ 0.20	0.787	0.885
ULO ↔ TLA	0.40	0.83	+ 0.43	0.787	0.885
Assessment ↔ TLA	0.80	0.73	−0.07	0.787	0.885

The interpretation of these patterns is addressed in the Discussion.

Discussion

This study addressed the question of how SOLO and Porter's Alignment Index can be integrated into a single protocol to quantitatively assess constructive alignment through curriculum artefact analysis. The demonstration using EDU101 shows that the integration is feasible and produces distinct, interpretable outputs that differentiate between pre and post curriculum conditions. The CAAP detected the alignment properties embedded in the designed unit, producing SOLO profiles and PAI values that reflect the known structural differences between the two curriculum designs.

The pre-intervention pattern

The pre-intervention outputs reveal a curriculum design with three structural problems that are each directly detectable through the CAAP. First, two of five ULOs use non-operational language, reducing the proportion of measurable outcomes to 0.60 and undermining the unit's stated cognitive ambition. This is precisely the pattern Brabrand and Dahl (2008) describe as characteristic of unaligned curricula, where the teacher's intention is obscured by outcome statements that cannot be operationalised or assessed. Second, the SOLO profile of the three operational ULOs is discontinuous, with representation at SOLO 2, 3, and 5 but no mid-level relational demand at SOLO 4, creating a gap in the progression from surface to deep learning. Third, the PAI values for ULO ↔ Assessment (0.50) and ULO ↔ TLA (0.40) both fall below the lower Fulmer (2011) critical value of 0.787, indicating alignment weaker than would be expected by chance. Examination of the proportion matrices reveals the source: the assessment operates entirely at SOLO 2–3 and does not test the SOLO 5 competence specified in ULO4, accessibility carries no weight in any artefact type, and teaching activity is uniformly concentrated at lower cognitive levels.

The pre-intervention Assessment ↔ TLA value of 0.80 warrants careful interpretation. This value does not indicate strong alignment in a pedagogically meaningful sense. Rather, it reflects structural consistency between two artefact types that both operate at SOLO 2–3 with similar topic coverage. As Kazemi (2022) notes, the PAI value alone does not reveal the nature or location of alignment, and in this case the underlying matrices reveal that the consistency is at a uniformly low level of cognitive demand. This illustrates the interpretive limitation identified by Fulmer and Polikoff (2014): a PAI value without matrix-level examination can be misleading.

The post-intervention pattern

Post-intervention, the SOLO profile shifts substantially. All four ULOs are operational, the SOLO 4 gap is filled with four verbs at the relational level, the SOLO average increases from 3.33 to 3.88, and the proportion of qualitative competences increases from 33.3% to 71.4%. These changes are directly reflected in the proportion matrices: the post-intervention ULO matrix distributes weight across all five topics and SOLO levels 3 through 5, and both the assessment and TLA matrices show corresponding shifts to higher cognitive demand.

The ULO ↔ TLA PAI of 0.83 post-intervention exceeds the lower Fulmer critical value of 0.787, indicating alignment stronger than would be expected by chance and representing a substantial increase from the pre-intervention value of 0.40. The ULO ↔ Assessment value of 0.70 represents meaningful improvement from 0.50 but does not yet reach the lower critical value threshold. The matrices suggest why: the redesigned ULOs contain multi-verb entries with SOLO 3 components that introduce lower-level proportions into the ULO matrix, while the assessment targets only SOLO 4–5. This residual tension is visible in the matrices and represents a natural focus for further redesign iteration.

The Assessment ↔ TLA decrease from 0.80 to 0.73 reflects reduced distributional overlap rather than a deterioration in alignment quality. Pre-intervention, both assessment and TLA operate at SOLO 2–3, producing high alignment through uniform low-level distributions. Post-intervention, assessment is concentrated at SOLO 4–5 while TLAs span SOLO 3–5, resulting in lower overlap and a reduced PAI. This pattern is consistent with a design in which teaching activities extend across levels to support higher-order assessment demands rather than mirror them exactly, illustrating Kazemi's (2022) point that matrices must be examined alongside PAI values to interpret alignment in context.

Contribution and implications

The demonstration establishes that the CAAP can detect and quantify the specific alignment properties it was designed to identify. As a designed artefact evaluated through a heuristic demonstration (Hevner et al., 2004; Siggelkow, 2007), the protocol's contribution is procedural: it provides a replicable, quantitatively grounded method for curriculum alignment analysis not previously available in the higher education quality assurance literature. Existing approaches rely primarily on qualitative audit tools and expert judgement (Ashwin et al., 2015), which do not produce quantitative indices or pre/post comparison capability. The development of digital tools to support evidence-based quality assurance in higher education has been identified as a significant practical need (Schellekens et al., 2023), and the CAAP addresses this need specifically for constructive alignment analysis at the unit level.

The availability of SOLO as the shared cognitive demand dimension resolves the key limitation of Porter's framework in higher education contexts identified by Kazemi (2022): the absence of a standardised, theoretically grounded cognitive taxonomy. By anchoring the column dimension of the Porter matrices in SOLO, the CAAP produces alignment indices directly interpretable in terms of the cognitive demand framework already in use in outcomes-based curriculum design, making the protocol practically usable by academic developers and curriculum leaders without specialist measurement expertise.

Limitations

The demonstration uses a single constructed unit designed to exhibit known alignment properties, and findings cannot be generalised to real curriculum contexts. The protocol's performance with ambiguous or inconsistently documented real-world artefacts has not been tested. The study uses a single coder, so classification decisions are not independently verified, though the codebook and calibration procedures provide transparency. The ordinal-to-interval assumption inherent in SOLO averaging is a recognised

simplification (Brabrand & Dahl, 2008), and PAI values do not support statistical significance testing beyond Fulmer (2011) benchmarks without supplementary methods (Fulmer & Polikoff, 2014). These limitations are appropriate to a protocol development study and point directly to the empirical research agenda the CAAP is designed to support.

Conclusion

This study demonstrates how the SOLO taxonomy can be integrated into Porter's Alignment Index to provide a unified quantitative method for measuring constructive alignment. The CAAP produces distinct and interpretable outputs that detect structural differences between curriculum designs and allow alignment to be examined at both the level of cognitive demand and artefact distribution. While demonstrated using a constructed case designed to exhibit known alignment properties, the protocol provides a transparent and replicable method for analysing alignment that extends beyond existing qualitative approaches.

The integration is achieved through a single conceptual move: adopting SOLO as the cognitive demand dimension of the Porter matrices. This creates a shared measurement language across ULOs, assessment tasks, and teaching activities, enabling alignment to be quantified rather than merely described. The CAAP operationalises this integration through a codebook, calibration procedure, and automated workbook that make the method accessible to practitioners without specialist measurement expertise.

The findings point to three directions for future research. First, empirical application of the CAAP to real curriculum artefacts is needed to test the protocol's performance beyond the controlled conditions of the demonstration case and to generate validity evidence across disciplines and institutional contexts. Second, multi-coder studies would allow inter-rater reliability to be established and the calibration procedures to be refined. Third, longitudinal application across curriculum review cycles would enable the protocol to contribute cumulative evidence about alignment patterns and their relationship to curriculum quality improvement over time.

The CAAP, SOLO Coding Reference Guide, workbook, and test dataset guide are openly available at [https://doi.org/10.17605/OSF.IO/RGTJS] to support replication, adaptation, and empirical extension of the method.

Declarations

Ethics statement

This study did not involve human participants, data collection from individuals, or access to sensitive materials. Ethical approval was not required.

Funding

This study received no external funding.

Conflicts of interest

The author declares no conflicts of interest.

Use of artificial intelligence

No artificial intelligence tools were used in the ideation, design, data generation, analysis, or interpretation of the research.

Data availability

The CAAP protocol, SOLO Coding Reference Guide, workbook, and test dataset guide are openly available (Sims, 2026).

CRediT author statement

Craig Sims: Conceptualisation, Methodology, Software, Validation, Formal Analysis, Investigation, Resources, Data Curation, Writing - Original Draft, Writing - Review and Editing, Visualisation.

References

1. Ashwin, P., Boud, D., Coate, K., Hallett, F., Keane, E., Krause, K., Leibowitz, B., MacLaren, I., McArthur, J., McCune, V., & Tooher, M. (2015). *Reflective teaching in higher education*. Bloomsbury Academic. https://www.bloomsbury.com/uk/reflective-teaching-in-higher-education-9781441197221/
2. Biggs, J. B. (1996). Enhancing teaching through constructive alignment. *Higher Education, 32*(3), 347–364. https://doi.org/10.1007/BF00138871
3. Biggs, J. B. (2003). *Teaching for quality learning at university* (2nd ed.). Open University Press.
4. Biggs, J. B., & Tang, C. (2007). *Teaching for quality learning at university* (3rd ed.). Open University Press.
5. Biggs, J. B., & Collis, K. F. (1982). *Evaluating the quality of learning: The SOLO taxonomy, structure of the observed learning outcome*. Academic Press.
6. Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal, 9*(2), 27–40. https://doi.org/10.3316/QRJ0902027
7. Brabrand, C., & Dahl, B. (2008). Constructive alignment and the SOLO taxonomy: A comparative study of university competences in computer science vs. mathematics. *Conferences in Research*

and Practice in Information Technology, 88, 3–17. https://vbn.aau.dk/ws/portalfiles/portal/212444193/CRPITV88Brabrand.pdf

8. Brabrand, C., & Dahl, B. (2009). Using the SOLO taxonomy to analyze competence progression of university science curricula. *Higher Education, 58*(4), 531–549. https://doi.org/10.1007/s10734-009-9210-4
9. Chan, C. C., Tsui, M. S., Chan, M. Y. C., & Hong, J. H. (2002). Applying the structure of the observed learning outcomes (SOLO) taxonomy on student's learning outcomes: An empirical study. *Assessment and Evaluation in Higher Education, 27*(6), 511–527. https://doi.org/10.1080/0260293022000020282
10. Fulmer, G. W. (2011). Estimating critical values for strength of alignment among curriculum, assessments, and instruction. *Journal of Educational and Behavioral Statistics, 36*(3), 381–402. https://doi.org/10.3102/1076998610381941
11. Fulmer, G. W., & Polikoff, M. S. (2014). Tests of alignment among assessment, standards, and instruction using generalized linear model regression. *Educational Assessment, Evaluation and Accountability, 26*(3), 225–240. https://doi.org/10.1007/s11092-014-9196-z
12. Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly, 28*(1), 75–105. https://doi.org/10.2307/25148625
13. Kazemi, J. (2022). *Measuring alignment between components of an educational system: An extension of Porter's alignment index* [Doctoral dissertation, City University of New York]. CUNY Academic Works. https://academicworks.cuny.edu/gc_etds/4983/
14. Porter, A. C. (2002). Measuring the content of instruction: Uses in research and practice. *Educational Researcher, 31*(7), 3–14. https://doi.org/10.3102/0013189X031007003
15. Porter, A. C., Smithson, J., Blank, R., & Zeidner, T. (2007). Alignment as a teacher variable. *Applied Measurement in Education, 20*(1), 27–51. https://doi.org/10.1080/08957340709336729
16. Schellekens, L. H., van der Schaaf, M. F., van der Vleuten, C. P., Prins, F. J., Wools, S., & Bok, H. G. (2023). Developing a digital application for quality assurance of assessment programmes in higher education. *Quality Assurance in Education, 31*(2), 346–366. https://doi.org/10.1108/QAE-03-2022-0066
17. Siggelkow, N. (2007). Persuasion with case studies. *Academy of Management Journal, 50*(1), 20–24. https://doi.org/10.5465/amj.2007.24160882
18. Sims, C. (2026). *Constructive Alignment Artefact Protocol (CAAP)* [Coding protocol and dataset]. OSF. https://doi.org/10.17605/OSF.IO/RGTJS
19. Stålne, K., Kjellström, S., & Utriainen, J. (2016). Assessing complexity in learning outcomes: A comparison between the SOLO taxonomy and the model of hierarchical complexity. *Assessment and Evaluation in Higher Education, 41*(7), 1033–1048. https://doi.org/10.1080/02602938.2015.1047319