International Journal of Science and Research Archive
eISSN: 2582-8185
Cross Ref DOI: 10.30574/ijsra
Journal homepage: https://ijsra.net/

(Research Article)

Check for updates

The Perceptual Invariance Framework: An Integrated Educational-Engineering Approach from Semantic Noise Elimination to Prompt Design in the Age of Artificial Intelligence

Nasip Demirkuş *

Department of Biology Education, Faculty of Education, Van Yüzüncü Yıl University, Tuşba, Van, Türkiye.

International Journal of Science and Research Archive, 2026, 19(03), 431-462

Publication history: Received on 02 May 2026; revised on 08 June 2026; accepted on 10 June 2026

Article DOI: https://doi.org/10.30574/ijsra.2026.19.3.1307

Abstract

This study integrates the core conceptual tools of the Perceptual Invariance Framework (PIF; also referred to in earlier work as the "Perceptual Invariance Theory"), developed by Demirkuş (2026), along two complementary axes. The first axis focuses on educational material design; it positions the phenomenon of "semantic entropy"—the systematic degradation of meaning during knowledge transmission—as one of the likely fundamental drivers of chronic educational failure, and offers three engineering solutions composed of the Generalization Rule, the Uniqueness Rule, and the EduCode Protocol. The second axis adapts PIF to human–machine educational communication established with generative artificial intelligence (GenAI); arguing that the semantic noise observed in educational materials re-emerges in communication with large language models (LLMs) in the form of hallucination and output variance, it proposes a three-layer "PIF-Prompt Model" for structured prompt design.

The study builds a theoretical bridge between Cognitive Load Theory (Sweller, 2024), Schema Theory (Anderson, 2020) and the communication model of Shannon and Weaver (1949), Bloom's (1968) Mastery Learning tradition, the 26 prompt principles of Bsharat et al. (2023), the chain-of-thought approach of Wei et al. (2022), and UNESCO's (2024a, 2024b) AI competency frameworks. When PISA 2022 (OECD, 2023) and Education Next's (2025) linguistic clarity analysis are considered together, a conceptual relationship is proposed whereby linguistic ambiguity may be a shared pathology in both human and machine-based educational communication; this relationship is a theoretical reading rather than causal evidence. In PIF—integrated with Bloom's Mastery Learning literature—the clarity standard is defined at two levels as initial assumptions requiring empirical calibration: the 95% rate is an applicable (operational) design target, while the 99% rate is positioned as the theoretical ideal to be approached.

It is explicitly acknowledged that the proposed Clarity-Indexed Scoring System and the EduCode Protocol require empirical validation; a multi-site Randomized Controlled Trial framework covering the 2025–2027 period ($N=2000$, $\alpha=0.05$, $\beta=0.80$, $d=0.3$) is presented to guide this validation process. PIF proposes to transform education from a probabilistic selection mechanism into a more predictable, measurable, and auditable engineering discipline in which the 95% applicable clarity target (99% theoretical ideal) is positioned as a design criterion. This study is a theoretical/model-developing work that presents testable hypotheses and implementation protocols.

Keywords: Perceptual Invariance Framework; PIF-Prompt Model; Educational Engineering; Cognitive Load Theory; Semantic Noise; Instructional Design; EduCode Protocol; Prompt Engineering; Generative Artificial Intelligence; Mastery Learning; Educational Equity

* Corresponding author: Nasip DEMİRKUŞ; ORCID: 0000-0003-4195-070X ·

1. Introduction

Education serves as the fundamental backbone for the continuity of civilization and the transmission of cumulative human knowledge. In practice, however, this transmission process is profoundly disrupted by semantic entropy (Demirkuş, 2024b). (The fundamental theoretical framework of this study was developed in Part I; references made in the text as "Demirkuş (2026)" point to this theoretical basis of the same integrated work. For the published citation, see Demirkuş, N. (2026, in press), WJAETS, 18(03), 081–097) When an analogy is drawn with the physical entropy of thermodynamics—which describes the tendency toward disorder in closed systems—semantic entropy in the educational context manifests as the gradual degradation of meaning from the moment information departs its source until it reaches the learner (Shannon and Weaver, 1949). It must be emphasized that this is an intuitive/explanatory analogy; whereas the increase of entropy in the second law of thermodynamics is inevitable in closed systems, semantic degradation in education is not inevitable and can be substantially reduced with well-designed materials.

This conventional perspective reflects a bias that ignores the systemic "noise" generated by the communication channel itself, placing the burden of error disproportionately on the student (Demirkuş, 2024b). Drawing on the Perceptual Invariance Framework (PIF) developed by Demirkuş, this study advocates a conceptual departure from probabilistic educational models; it advances the idea of transforming education from a practice based on chance and individual interpretation into a more predictable, measurable, and auditable engineering discipline.

This thesis converges conceptually with the linguistic clarity analysis conducted by Education Next (2025): when 7 linguistic clarity criteria derived from more than 330 peer-reviewed studies were applied, no material of a top-ranked mathematics program achieved a full score. The problems include complex conditional statements, inconsistent terminology, and references requiring cultural background knowledge; these cause the student to expend cognitive capacity on parsing language rather than on mathematical thinking.

The core hypothesis asserts the following: the value of any educational element should be evaluated not by the difficulty level or discriminatory power emphasized in classical psychometrics, but by its capacity to elicit a single, stable, and consistently accurate mental response in an extremely high proportion of the target audience (applicable target ≥ 95%; theoretical ideal 99%).

1.1. The Semantic Noise Crisis in Education and Communication Pathology

The most critical pathology confronting contemporary education systems is the phenomenon defined in communication sciences as "Semantic Noise." According to the foundational Communication Model of Shannon and Weaver (1949), noise refers to any distortion that occurs between the encoding of a message at the source and its decoding at the destination. Semantic noise refers to coding errors in which the language, symbols, technical jargon, visuals, or cultural codes used by the sender trigger a meaning construction in the recipient's mind that is fundamentally different from what the sender intended (Crystal, 2019).

1.1.1. The Disconnect Between Intent and Perception

For example, in a physics context, the use of the word "work" may, based on everyday usage, evoke associations such as "occupation," "labor," or "effort" in the student's mind. The teacher, however, intends the specific mathematical definition $W = F \times d$. This semantic mismatch leads to a loss of meaning even though the acoustic signal is technically transmitted successfully.

1.1.2. Cognitive Automaticity

To prevent this chaotic variance, the Perceptual Invariance Framework aims to process core concepts at the level of cognitive automaticity (Demirkuş, 2024a; Shiffrin and Schneider, 1977; Logan, 1988). The intent here is that a properly engineered educational material should trigger correct information retrieval in a fast, stable, and low-effort manner. This phenomenon—expressed in earlier work through the "mental reflex" analogy—differs from the unconscious physiological reflex at the spinal-cord level; the speed and stability in question are a cognitive process and are more accurately explained by the cognitive automaticity literature. Hence the term "mental reflex" should be understood not as a physiological identity but as an analogy describing automatized retrieval.

This approach converges with research on perceptual learning and the development of expertise. Studies show that experts (e.g., radiologists, chess masters) differ from novices in their ability to extract "invariants" from their environment rapidly and holistically (Gibson, 1969; Goldstone, 1998). Although Dewey (1910) and Schön (1983) foregrounded "reflective thinking," the framework argues that higher-order thinking is possible only when basic concepts are processed with reflexive speed and accuracy (Sweller, 2024).

1.2. Analysis of the Current Situation through Global Data: Mental Tissue Mismatch

One of the fundamental causes of failure in education is "Mental Tissue Mismatch," which denotes the deep chasm between the intended curriculum and the attained curriculum (Demirkuş, 2024a). The possible effect of semantic noise can be observed indirectly in international statistics. An important caveat is required: the indicators presented in the table below are multifactorial phenomena; student failure cannot be reduced to semantic noise alone. Socioeconomic status, teacher quality, curriculum design, school facilities, and language policies are also determining factors. Semantic noise should be regarded as only one component of this multi-component picture, and the relationship established with the data is explanatory rather than causal.

Table 1 Global and National Indicators of Semantic Difficulties in Education

Data Source	Indicator	Inference and Analysis
UNESCO GEM Report (2021–2024)	Approximately 40% of students globally cannot understand materials due to linguistic complexity or contextual gaps.	About half of the global student population encounters a semantic noise barrier. This deficiency is a design problem, not an intelligence problem.
PISA 2022 (Türkiye)	29% of students perform below Level 2 in reading skills (OECD, 2023).	This rate is equivalent to the OECD average (29%); however, students can 'decode' the text yet fail at semantic reconstruction.
OECD Average (PISA)	45% of students struggle to interpret complex texts as the author intended (OECD, 2023).	The intended message is lost in transmission for nearly half the group.
PISA 2022 Mathematics (Türkiye)	Score: 453 (OECD avg.: 472). The critical failure is in 'converting context into mathematics' (OECD, 2023).	The problem is not mathematical competence but comprehension of mathematical language. Semantic noise in the item text blocks schema access.
Education Next (2025)	No material of a top-ranked mathematics program fully met the 7 linguistic clarity criteria derived from 330+ peer-reviewed studies.	Complex conditionals, inconsistent terminology, and cultural references increase cognitive load, forcing students to parse language rather than do mathematics.
World Bank (2022–2024)	'Learning Poverty' in developing countries rose to 70% after the pandemic.	The vast majority of 10-year-olds cannot understand simple text. A systemic inability to build basic semantic bridges is at issue.

Note. Compiled by the author by synthesizing data from OECD (2023), UNESCO (2021), the World Bank (2022, 2024), and Education Next (2025).

1.3. The Clarity Principle, the "White Paper" Test, and Context-Independent Invariants

The Perceptual Invariance Framework sets out the Clarity Principle as a quality standard. This principle is defined at two levels: the formation of a single, consistent semantic equivalent by 95% of the target audience is the applicable (operational) design target; 99% is positioned as the theoretical ideal to be approached. To draw an engineering analogy, just as a load-bearing structure is expected to remain standing with high reliability, educational definitions are expected to bear the weight of meaning with high precision. It must be emphasized that these numerical values are initial criteria requiring empirical calibration.

The Clarity Principle is conceptually linked to Bloom's (1968) Mastery Learning tradition: a meta-analysis of 36 mastery learning studies reported an average effect size of d=0.59 (Winget and Persky, 2022). PIF's high clarity standard can be interpreted as a material-focused conceptual extension of this empirical tradition. While Bloom foregrounds the student's re-study, PIF foregrounds the re-engineering of the material; both approaches share a 'teaching everyone' philosophy. Nevertheless, it must be noted that the empirical effect of mastery learning does not directly prove PIF's specific clarity thresholds (95%–99%); these thresholds are targets that must be tested through pilot studies (Bloom, 1968; Guskey, 2024).

1.3.1. The White Paper Test Analogy

When an educator shows the class a white sheet of paper and asks "What color is this?" and the whole class answers "White," "Pure White," or "Off-White," these answers indicate a perceptual invariance rate of 99% (Demirkuş, 2024a). Variability in student responses should reflect the student's knowledge state, not the ambiguity of the stimulus.

1.3.2. Invariants with a Stable Core Meaning (Context-Resistant)

The most extreme example of the stability sought in education is invariants with a highly stable (context-resistant) core meaning: propositions that largely preserve the same meaning across cultural, social, or individual differences. For example, a proposition such as "a non-living entity does not produce a biological response" has the quality of a stable constant closed to interpretation. Educational curricula should be built primarily not upon interpretable, ambiguous, or "artistic" expressions, but upon such context-resistant invariants. This approach is directly connected to J. J. Gibson's Perceptual Learning Theory (Gibson, 1969) and to Felix Klein's Erlangen Program (Klein, 1872)—the study of structures that remain unchanged under geometric transformations.

2. The Clarity Dimension in Psychometric Assessment: A New Perspective on Measurement and Evaluation

The most striking intervention of the Perceptual Invariance Framework manifests in the domain of measurement and evaluation. Traditional psychometrics, particularly Classical Test Theory (CTT), evaluates the quality of a test item according to "Discrimination" and "Difficulty" indices. Perceptual Invariance fundamentally questions these parameters and redefines the goal of education not as "selection/elimination" but as "teaching/acquisition" (Demirkuş, 2024a).

2.1. A Paradigm Shift from Classical Test Theory (CTT)

In the standard educational measurement literature, item statistics are defined as follows (Turgut and Baykul, 2019):

- **Item Difficulty Index (pj):** The ratio of correct answers to the total number of respondents (between 0.00 and 1.00). CTT argues that an average difficulty of 0.50 is ideal for maximizing variance and reliability. An item that half the class fails is considered "good" because it spreads scores, allowing ranking.
- **Item Discrimination Index (rjx):** Indicates how well an item distinguishes between high- and low-performing students. An item answered correctly by everyone (pj=1.00) has a discrimination value of 0.00 and is discarded as "non-functional" in selection examinations.

PIF's Objection: When the goal is mastery learning, CTT falls short. When 100% of students answer an item correctly (pj=1.00), CTT regards this as the item's failure to discriminate. Perceptual Invariance, by contrast, accepts it as the indicator of a perfect educational material. This transformation brings a philosophical conflict to the surface: does the education system exist to race students against one another (ranking), or to verify whether they have attained specific learning objectives?

Table 2 Comparative Analysis of Perceptual Invariance and Traditional Methods

Feature	Traditional Approach (CTT)	Perceptual Invariance Approach (PIF)
Primary Goal	Ranking and Eliminating Students	Teaching Everyone (≥ 95% Success)
Value Criterion	Difficulty and Discrimination	Clarity and Comprehensibility
Ideal Rate	50% Success (p=0.50)	≥ 95% Success (p ≥ 0.95)
Interpretation of Low Success	"Student doesn't know" or "Item is hard"	"Material is defective" or "Semantic Noise exists"
Scoring Logic	Fixed points (usually)	Weighted points based on the Clarity Index
Source of Error	Student Ignorance / Chance	Material Ambiguity / Perceptual Deviation
Analogy	Bell Curve (Competition)	Barcode Reader (Verification)
Historical Comparison	Norm-referenced standardized tests	Bloom (1968) Mastery Learning + PIF

Note. Compiled by synthesizing Turgut and Baykul (2019), Demirkuş (2024a), and Bloom (1968).

Important note: The comparison below is not an "all-or-nothing" choice but a difference in emphasis. PIF does not exclude CTT's difficulty and discrimination parameters; these parameters are indispensable in norm-referenced, highstakes selection-ranking examinations. PIF should be positioned as a complementary framework that adds a "clarity" dimension to these parameters in the context of formative assessment and material development.

2.2. Comparison of PIF with Bloom's Mastery Learning

PIF's critique of Classical Test Theory has deep roots in the history of education. Bloom (1968) argued that standard norm-referenced tests should be used not for "selecting and ranking" but for "learning and attaining mastery," and built the mastery learning model on this philosophy. PIF takes this philosophy one step further: it asks not only "Did the student reach the mastery threshold?" but also "Was the material clear enough to make this possible?"

In the Mastery Learning model, assessments cease to be outcome-oriented tools and become part of the learning process; students cannot proceed to the next unit before reaching a certain mastery threshold. A meta-analysis of 36 mastery learning studies reported an average effect size of d=0.59 (Winget and Persky, 2022). This provides indirect support that mastery-based approaches produce a meaningful effect on learning; however, it must be emphasized that this meta-analysis did not test PIF-specific material engineering, and that the 15–25% comprehension increase predicted by PIF is not a proven result but a hypothesis to be tested. In this context, the fundamental difference between Mastery Learning and PIF is the following: whereas Bloom's model structures corrective learning activities around the student's learning process (aiming to bring everyone to success through additional time and corrective instruction), PIF treats ambiguity in material design as a more central problem and shifts a significant portion of the responsibility onto the re-engineering of the material. This difference in emphasis is also ethically critical; nevertheless, it must be noted that the two approaches are not mutually exclusive but complementary.

Table 3 Comparison of PIF with Bloom's Mastery Learning Model

Dimension	Bloom's Mastery Learning (1968)	PIF — Perceptual Invariance (Demirkuş, 2026)
Primary Goal	Reaching a predetermined mastery threshold	Formation of an identical semantic equivalent in ≥ 99% of the target audience
Assessment Function	Formative assessment: a means of the learning process	Clarity-Indexed Scoring: measures material quality
Interpretation of Failure	Student has not yet reached the mastery threshold → additional time and corrective instruction	Material is defective → the source of semantic noise must be identified
Standard	Predetermined proficiency threshold (usually 80%)	≥ 99% Perceptual Invariance Rate
Effect Size (Evidence)	Meta-analysis: d=0.59 (36 studies)	No empirical validation yet; RCT proposed (N=2000)
Fundamental Difference	Student-centered: the student's re-study activity	Material-centered: re-engineering of the material

Note. Compiled as a synthesis of Bloom (1968), Winget and Persky (2022), and Demirkuş (2026).

2.3. Revised Scoring Methodology: The Clarity-Indexed System

The Perceptual Invariance Framework proposes a new Scoring Methodology that determines the "market value" of an item. In this system, an item's point value is directly proportional not to its difficulty but to its "Clarity"—the rate at which it is understood identically by everyone.

Score = Base × [(w_1 × p) + (w_2 × r)]

(Here, the weights w_1=0.7 and w_2=0.3 together with the ≥ 95% target are not precise criteria but initial assumptions that must be calibrated through pilot studies.) The clarity criterion should not be reduced merely to the "rate of correct understanding"; a valid Clarity Index should be evaluated together with the content and construct validity of the measurement, its reliability (internal consistency), inter-rater agreement (e.g., Cohen's κ or ICC), and a reduction in the variance of student responses.

- w_1 (Clarity Weight) = 0.7
- w_2 (Discrimination Weight) = 0.3

- p = Comprehension Rate (the percentage of people who correctly understood the item)
- r = Discrimination Index

Nevertheless, one point must be underlined: the weights w_1=0.7 and w_2=0.3 rest on the theoretical rationale that 'clarity should be prioritized over discrimination' and may require calibration in different pedagogical contexts. This uncertainty is a limitation explicitly acknowledged in the article; the parameters must be recalibrated through inter-rater reliability studies.

2.4. Example Application

Base Score: 10 · Comprehension Rate (p): 0.95 · Discrimination (r): 0.60

Score = 10 × [(0.7 × 0.95) + (0.3 × 0.60)] = 10 × 0.845 = 8.45

This approach converges with Differential Item Functioning (DIF) analysis, which removes bias. **Central Bank Analogy:** just as a banknote (item) is tested before entering circulation, an item should not be released for use until it is confirmed that 95% of students 'understand it correctly.' Justice is sought not in "equality of difficulty" but in "certainty of clarity" (Demirkuş, 2024b).

3. Educational Material Design Engineering: The Generalization and Uniqueness Rules

Achieving high invariance rates requires transforming material preparation from an art form into an engineering discipline governed by strict rules. The two pillars of this design are the Generalization Rule and the Uniqueness Rule. These are deeply integrated with Cognitive Load Theory (Sweller, 2024) and Schema Theory (Anderson, 2020).

3.1. The Generalization Rule: Establishing the Cognitive Framework and Schema Theory

The Generalization Rule mandates that, before moving on to details, the most inclusive, indisputable, and "fingerprint-like" definition of a subject must be established to delimit its boundaries.

- **Objective:** To create a "Common Denominator" in the entire target audience; to eliminate "conceptual confusion" before it begins.
- **Cognitive Mechanism:** According to Schema Theory (Anderson, 2020; Bartlett, 1932), new information acquires meaning only when associated with existing structures. Generalization offers the mind a simple, inclusive schema, allowing new information to settle into this framework.
- **Relation to Cognitive Load:** Sweller (2024) shows that reducing "element interactivity" is the key to managing cognitive load. A generalized definition functions as a "chunking" mechanism that reduces element interactivity by grouping complex details under a single, invariant conceptual label.
- **Application Example (The Cell):** When introducing the cell in biology, the Generalization Rule should be applied first: "The cell is the fundamental unit of living things in terms of structure, function, and hereditary organization." This definition is an invariant roof covering everything from an ostrich egg to a bacterium and, like an "advance organizer" (Ausubel, 1960), creates a stable mental anchor. It should be noted that in Ausubel's original definition advance organizers are relatively abstract, high-level frameworks; the generalization definition here, while functioning as an advance organizer by providing a high-level, inclusive roof, is not strictly identical to a classic advance organizer because it also contains concrete criteria.

3.2. The Uniqueness Rule: Decomposition and Differentiation of Content

Once the general roof is established, the Uniqueness Rule comes into play, defining how each variation under this roof is distinguished from the others.

- **Objective:** To prevent the confusion of types (Unique Shares) within the general concept.
- **Differentiation:** This reflects the "differentiation" in Gibson's (1969) perceptual learning process. Starting from the whole (Generalization), the learner learns to perceive the fine distinctive features (Uniqueness).

3.2.1. Application Example (The Pen)

- **Generalization:** "A tool used for writing on suitable surfaces." (Covers all pens.)
- **Uniqueness → Pencil:** Graphite core, erasable.
- **Uniqueness → Fountain Pen:** Ink reservoir, liquid flow.

Each type is defined by its "Unique DNA Barcode"; this prevents conceptual interference known in memory psychology as Proactive or Retroactive Inhibition.

3.3. Integration with Gestalt Perception Principles

The Perceptual Invariance Framework mandates the conscious use of Gestalt Psychology principles in material design (Demirkuş, 2024a). The human mind is structured to seek order and wholeness; educational engineering must serve these tendencies.

Table 4 Application of Gestalt Principles in Educational Material Design

Gestalt Principle	Problem in Education (Noise)	Perceptual Invariance Solution
Figure–Ground	The student cannot distinguish the main idea from the details.	The main concept (Generalization) is foregrounded; the details (Ground) are simplified. Visual clutter is eliminated.
Closure	Missing information is completed with bias.	Definitions are given with 'fingerprint' precision; no room is left for subjective interpretation.
Similarity	Different concepts are confused as similar.	The Uniqueness Rule clarifies differences by highlighting distinctive features (e.g., Virus — Bacterium).
Prägnanz (Simplicity)	Complex narration increases cognitive load.	The simplest 'common denominator' definition is given first to establish a stable cognitive anchor.

Note. A synthesis of Demirkuş (2024a) and Mayer (2009).

4. Linguistic Barriers, Semantic Noise, and Conceptual Misconceptions

A critical issue addressed in this study is the structure of natural languages, which are far from scientific precision. Linguists such as Crystal (2019) and Pinker (1994) emphasize that languages evolved not for scientific exactitude but for general communication. These languages are replete with polysemy, homonymy, and idiomatic irregularities that generate high semantic entropy.

4.1. The Mental Pool and Coding Mismatch

There is a deep chasm between the human "Inner World" (Mental Pool) and the "Outer World" (Linguistic Expression). When a Japanese, Turkish, or German person looks at a "Red Apple," they experience a highly similar biological and imagistic response in their visual cortex. Human hardware is common, and perceptual invariants are detected similarly (Demirkuş, 2024a). Yet chaos ensues when one tries to express this image: one says "Apple," another "Elma," another "Pomme." These are arbitrary labels that do not reflect the nature of the object.

Conceptual Misconceptions: while "work" means "effort" in everyday language, in physics it corresponds to the formula $W = F \times d$. When a student is told, "I pushed the wall and got tired but did no work," natural language (getting tired = work) conflicts with scientific language (no motion = no work), breaking invariance.

4.2. Solution: The Mathematical Barcode Language and the EduCode Protocol

The ultimate solution is to transform languages into a "Mathematical Barcode Language." This approach is rooted in Leibniz's Characteristica Universalis (1666), a vision of an alphabet that expresses all human thought through mathematical symbols (Demirkuş, 2024a). Its primary and most urgent application targets the STEM fields, where conceptual precision is paramount.

EduCode converges with the Controlled Natural Language (CNL) tradition (Kuhn, 2014). A CNL is a subset of natural language that restricts the permitted syntax and vocabulary in order to reduce ambiguity. SNOMED CT in medical informatics, CAS numbers in chemistry, and UniProtKB in bioinformatics are applications of this logic in different fields. EduCode is not an invention from scratch; it is the adaptation to the educational context of controlled terminology systems—such as SNOMED CT in medical informatics and UniProtKB in bioinformatics—that have operated successfully for decades; it is the transfer of a proven ontology-engineering paradigm to the learning sciences (Kuhn, 2014).

EduCode Protocol Example (JSON Structure)

{ "concept_id": "BIO.007.2024",

"universal_name": "Photosynthesis",

"mathematical_representation": "$6CO_2 + 6H_2O \rightarrow C_6H_{12}O_6 + 6O_2$",

"related_concepts": ["cellular_respiration", "chlorophyll"],

"difficulty_level": 2,

"prerequisites": ["BIO.001.2024"],

"multilingual_labels": {

"tr": "Fotosentez",

"en": "Photosynthesis",

"es": "Fotosíntesis",

"de": "Photosynthese" }

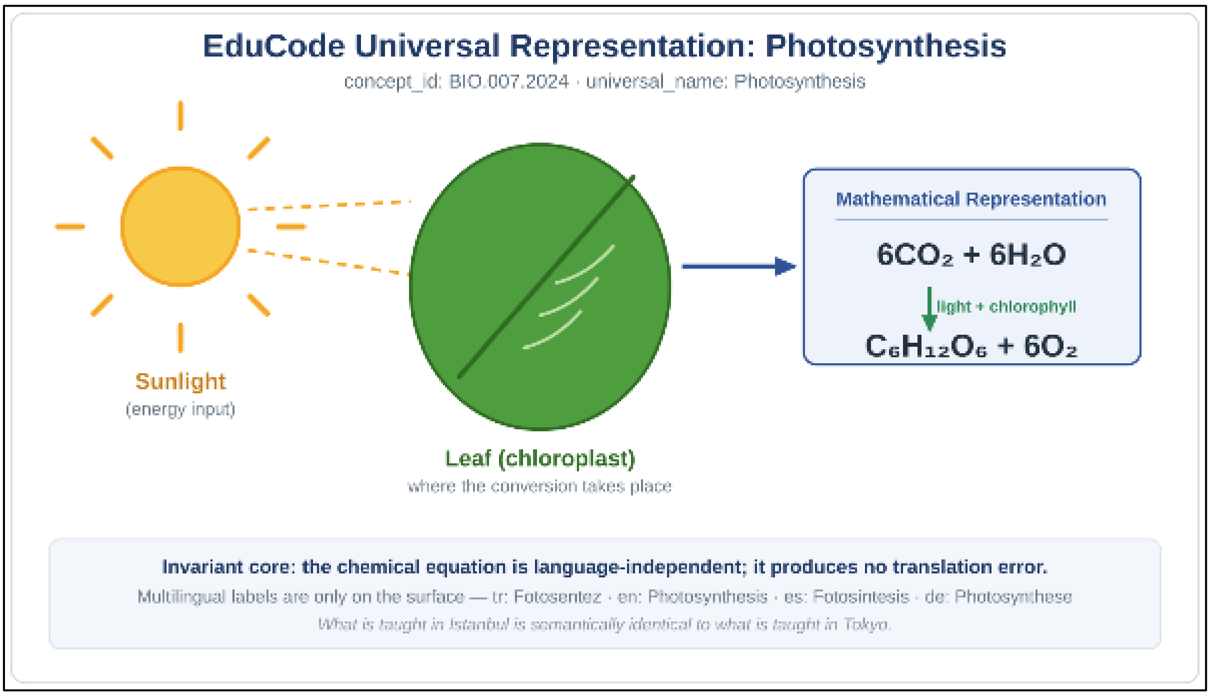

Figure 1 The EduCode Transformation Process — the invariant representation of the concept "Photosynthesis" via the sun (energy input), the leaf (chloroplast), and the language-independent mathematical equation ($6CO_2 + 6H_2O \rightarrow C_6H_{12}O_6 + 6O_2$)

4.2.1. Protocol Analysis:

- **Universal Identity:** "BIO.007.2024" points to the same biological reality on a global scale; it functions as the primary key in a global knowledge database.
- **Mathematical Representation:** Chemical formulas offer a precision free from translation errors.
- **Multilingual Labels:** The concept rests on the code; "Photosynthesis" becomes merely a label. This structure ensures that what is taught in Istanbul is semantically identical to what is taught in Tokyo.

Three-Stage Transition Plan

- **Stage 1 (2025–2027):** Widespread use of standard symbols in STEM; adding QR codes to digital materials. These QR codes link to EduCode definitions, providing an immutable reference point.
- **Stage 2 (2028–2030):** Creation of an international "Educational Concept Library" (EduOntology). Clicking a symbol reveals the invariant definition in all languages; the platform functions as the "Central Bank" that regulates the semantic integrity of concepts.
- **Stage 3 (2030+):** Combining natural language and mathematical symbols in a hybrid system; realizing Leibniz's vision of a universal language.

4.2.2. Governance and Feasibility of EduCode

For the EduCode Protocol to move from an abstract proposal to an applicable infrastructure, a governance model is required. In the proposed framework, responsibilities may be distributed as follows: (a) the "concept registration" process is conducted by a national education authority (e.g., the Board of Education in the Turkish context) together with disciplinary committees of field experts; (b) each concept code is approved after review by at least two independent field experts and a clarity audit by a linguistics/measurement expert; (c) erroneous or contested codes are corrected through version control and a feedback mechanism, with each code's change history kept traceable; (d) the system is built on open-access and interoperable standards. The central registry, peer approval, and versioning practices observed in established ontologies such as SNOMED CT and UniProtKB can be taken as direct references for this model (Kuhn, 2014).

Adaptation beyond STEM should be handled by grading concepts according to their degree of precision. In fields that cannot be represented by a mathematical formula (history, literature, social sciences), EduCode can be adapted not as a full mathematical barcode but as a structured record containing descriptive metadata, a relational network (ontological links), and contextual boundary conditions. In these fields the goal is not to eliminate interpretive diversity but to clear the inclusive core (common denominator) of the definition of ambiguity while explicitly marking the interpretive layer. Thus EduCode can be positioned as a cross-disciplinary, scalable framework that distinguishes the precision-requiring core from the interpretation-open layer. The validity of these adaptations must be tested through pilot studies.

4.3. The Relationship between PIF and EduCode: A Framework of Meaning and Its Coded Application

At this point, the conceptual relationship between PIF and EduCode must be stated explicitly, because the two components do not lie on the same plane of abstraction. PIF is a framework, while EduCode is the coding system in which this framework is applied to course and concept design. In other words, PIF answers the question "How is meaning made clear, invariant, and unmistakable?" while EduCode answers "How is this clear and invariant meaning coded, stored, and reused in educational materials?"

PIF's conceptual tools developed in the previous sections—reduction of semantic noise, the Generalization Rule, the Uniqueness Rule, and Clarity-Indexed Verification—translate into direct operational counterparts on the EduCode side. EduCode fixes each concept with a standard code, provides the concept's common-denominator (generalization) definition, specifies its distinctive features, and makes it measurable whether the student has understood the concept correctly. Thus perceptual invariance—a theoretical goal whereby the same concept is understood the same way by different people—becomes an applicable criterion at the coding level. Table 5 summarizes PIF principles and their EduCode counterparts.

Table 5 Mapping PIF Principles to Their EduCode Counterparts

PIF Principle (Framework)	EduCode Counterpart (Application)
Reduces semantic noise	Fixes the concept with a standard code
Generalization Rule	Provides the concept's common-denominator (core) definition
Uniqueness Rule	Specifies the concept's distinctive ("fingerprint") features
Clarity-Indexed Verification	Measures whether the student has understood the concept correctly
Perceptual invariance goal	Ensures the same concept is understood the same way by different people

Note. Compiled on the basis of the author's conceptual schema; PIF's theoretical principles are mapped to EduCode's operational counterparts.

4.3.1. Application Example: The Concept of "Photosynthesis"

From the PIF perspective, the goal in the photosynthesis example is for students to grasp the concept not in divergent and erroneous ways but with a single, stable, and shared core meaning. From the EduCode perspective, the same concept is transformed into a language-independent, reusable concept record:

- **Concept Code:** BIO-BOT-FOTO-001 (corresponds to the same entity as the numeric record BIO.007.2024 shown in §4.2; one is a human-readable hierarchical label, the other a numeric primary key).
- **Universal Name:** Photosynthesis
- **Generalization (common denominator) Definition:** The production of organic nutrients by converting light energy into chemical energy.
- **Chemical (mathematical) Representation:** $6CO_2 + 6H_2O \rightarrow C_6H_{12}O_6 + 6O_2$
- **Related Concepts:** chloroplast, chlorophyll, light, glucose, oxygen.
- **Prerequisite Knowledge:** cell, organelle, energy, matter cycle.

In the briefest terms, PIF is the engineering of meaning; EduCode is the coded application of this engineering of meaning in educational materials. This distinction is also decisive for the PIF-Prompt Model developed in Part II: the framework (PIF) specifies which semantic qualities a prompt must possess, while the coding system (EduCode) sets out how these qualities are templated in a machine-readable and reusable form.

5. Proposed Research Model: Methodological Framework

A comprehensive empirical research model is proposed to rigorously validate the theoretical claims of Perceptual Invariance. This model is designed to compare the effectiveness of "engineered" materials with traditional "artistic" materials.

5.1. Enhanced Pilot Study Design

Hypothesis: Educational materials designed in strict accordance with Perceptual Invariance principles (Generalization + Uniqueness + EduCode) will exhibit statistically significant improvements in comprehension fidelity, retention rates, and cognitive efficiency compared to traditional materials.

5.1.1. Sample Justification (Power Analysis)

- **Total Participants:** N=2000 (based on a power analysis with α=0.05, β=0.80, effect size d=0.3). The N=500 figure in the first version of the abstract has been corrected to N=2000 in line with this rationale. A clarification is required here: for a two-group independent-samples comparison at a small effect size of d=0.3, the minimum required sample is approximately N≈350 (≈175 per group; G*Power). The N=2000 value is far above this minimum and was chosen so that not only the basic comparison but also stratified subgroup analyses across four educational levels (primary, middle, high school, university) can be conducted with adequate statistical power. At the pilot stage, the power analysis will be recomputed based on observed effect sizes.
- **Group 1 — Experimental:** Perceptual Invariance materials (n=1000).
- **Group 2 — Control:** Traditional materials (n=1000).
- **Stratification:** Distribution across primary (400), middle (400), high school (600), and university (600) levels.
- **Control Variables:** Prior knowledge (pre-test), socioeconomic status, learning styles, cultural background, teacher quality.

5.1.2. Advanced Measurement Tools

- **Comprehension Test:** 20 multiple-choice items with DIF analysis; evaluated using the Clarity-Indexed Scoring system.
- **Eye-Tracking Analysis:** Used to measure fixation and saccade patterns. Hypothesis: students using PIF materials will show more focused fixation behavior on core concepts and fewer regressive saccades.
- **EEG/fNIRS:** Used to measure cognitive load via prefrontal cortex oxygenation and theta/beta wave ratios.
- **Delayed Recall Tests:** Administered 1 week, 1 month, and 6 months after instruction to measure schema durability.

5.2. Expected Outcomes

- **Comprehension:** Theoretically predicted (at the hypothesis level): a 15–25% increase in comprehension rates in the PIF group.
- **Cognitive Load:** A 25–40% reduction in cognitive load measures.
- **Retention:** A 30–50% improvement in 6-month delayed recall.
- **Equity:** Reduced performance gaps across demographic subgroups.

Note. The expected outcomes above are theoretical upper-bound estimates and are not supported by any pilot data. In particular, a 30–50% improvement in delayed recall is a magnitude rarely seen in the educational-intervention literature; such preliminary estimates carry the risk of effect-size inflation. Therefore these ranges should be read not as evidence but as testable hypotheses, open to downward correction through pilot studies.

5.3. Longitudinal Cohort Study Design

- **Duration:** A 3-year longitudinal study conducted on the same cohort.
- **Measures:** Annual assessments of conceptual understanding, problem-solving ability, and motivation.
- **Comparison:** Schools applying the traditional curriculum versus schools applying Perceptual Invariance.
- **Outcomes:** Long-term educational trajectories, STEM career choices, and lifelong learning patterns.

Part II — application: perceptual invariance and prompt design in the age of artificial intelligence

The PIF-Prompt Model for Reducing Variance in Human–AI Communication

---

6. Pedagogical Transformation in the Age of Artificial Intelligence

The proliferation of generative artificial intelligence technologies is structurally reshaping the field of education. Large language models such as ChatGPT, Claude, and Gemini have become embedded in the core of processes such as content generation, feedback provision, item preparation, and assessment (Qian, 2025). Yet the pedagogical value of these systems depends less on the power of the model than on the quality of the instruction (prompt) the user produces. Lee and Palmer (2025) show that the quality of prompt design is the most fundamental guiding mechanism determining the quality of the outputs obtained from AI. This has turned prompt engineering from a merely technical skill into an interdisciplinary field at the intersection of linguistics, cognitive psychology, and instructional design (Anowar and Khatun, 2026).

From the perspective of the educational sciences, this paradigm shift raises a theoretical question: is there a functional identity between the semantic noise that arises in a teacher's in-class explanation and the ambiguous prompt a user sends to an LLM? The Perceptual Invariance Framework developed in Part I locates the primary pathogen of educational failure not in student deficits but in the structural flaws of the communication channel. Transferring the same conceptual framework to LLM interactions shows that prompt ambiguity re-manifests in the form of hallucination, output variance, and task drift (Lakera, 2026; Meincke et al., 2025).

This section advances three core claims. **First claim:** natural language inherently carries structural ambiguities (linguistic multi-interpretability, in the sense that the same expression is decoded differently in different contexts and minds) as a transmission channel for LLMs. **Second claim:** PIF's Generalization and Uniqueness rules converge with the context-provision and specificity principles in the prompt-engineering literature (Bsharat et al., 2023); this convergence is not a coincidence but arises from the same root of communication pathology. The added value of the PIF-Prompt Model is not to rename these principles, but to ground them in a pedagogical framework (perceptual invariance), to define a measurable verification criterion via the clarity index, and to make the reduction of semantic noise an explicit design goal. **Third claim:** PIF's 99% Clarity Index offers a new and operationalizable quality criterion for the pedagogical validity of LLM outputs.

These three claims can be concretized by an in-class observation conducted by the author in a teaching context: when the same ambiguous prompt — for example, "Write me something about photosynthesis" — is given to four different large language models (ChatGPT, Gemini, Copilot, and Claude), the models produce outputs that differ markedly in pedagogical level, format, and cognitive load. This is not a controlled experiment but an illustrative/anecdotal observation describing the phenomenon of output variance; nevertheless, the inconsistent decoding of the same weak prompt across models exemplifies, at an applied level, PIF's claim that semantic noise re-manifests in human–machine communication in the form of output variance and hallucination.

7. Theoretical Bridge: PIF ↔ Prompt Engineering

7.1. Cognitive Load Theory and Human–AI Interaction

With Sweller's (2024) recent revision, Cognitive Load Theory emphasizes that the limited capacity of working memory is the fundamental constraint on learning. The theory defines three types of load: intrinsic load inherent to the concept, extraneous load arising from the presentation of the material, and germane load devoted to schema construction. This triad has a direct counterpart in the prompt-engineering context.

The cognitive offloading studies of Stadler et al. (2024) and Risko and Gilbert (2016) show that GenAI use produces a paradoxical effect: while models reduce cognitive load on the surface, poorly designed prompts create a dependency pattern that weakens deep thinking. From the student's standpoint the problem is this: when a prompt is sufficiently ambiguous, the GenAI tool produces an "approximate answer" and the student adopts it without constructing a schema. This is the exact opposite of PIF's concept of cognitive automaticity ("mental reflex"); it feeds epistemic laziness instead of automatized, stable retrieval. To mitigate this risk, the PIF-Prompt Model prescribes guiding mechanisms in the Uniqueness Layer such as desired/undesired output examples and post-output verification (e.g., asking for a functional rather than a rote definition and comparing the answer with the expected core definition). Neuro-cognitive validation (fNIRS- or EEG-based) studies in this area remain an important research gap.

Nevertheless, reducing cognitive load under all conditions is not necessarily beneficial for learning. Bjork and Bjork's (2011) framework of "desirable difficulties" shows that a certain productive effort—such as retrieval difficulty, spaced repetition, and variability—serves a germane function in consolidating knowledge in long-term memory. Therefore the noise elimination PIF aims at should not turn into a "semantic sterility" that also erases the productive difficulty that nourishes learning. The critical distinction is this: what PIF aims to eliminate is the extraneous/semantic noise arising from a coding error between intent and perception; the germane load arising from cognitively wrestling with a concept, making inferences, and building one's own schema must be preserved—indeed encouraged through design. Accordingly, while the Generalization Layer aims to lower intrinsic load, the Uniqueness and Verification layers aim to preserve productive effort by placing the student in the position of an active auditor rather than a passive consumer (cf. §11.2). Empirically determining the optimal balance between noise reduction and the preservation of desirable difficulty—especially via delayed-recall measures—remains an open research question.

7.2. The Shannon–Weaver Model and the Transfer of Semantic Noise to the Machine

Shannon and Weaver's (1949) communication theory established that a message transmitted from a source to a receiver is exposed to various types of noise, the most critical being semantic noise arising from inconsistencies in linguistic and symbolic coding (Crystal, 2019). As Laixi (2025) sets out, the inner workings of LLMs rest on probabilistic pattern completion over tokens; models respond to patterns rather than "understanding" intent. The controlled experiments conducted by Bsharat et al. (2023) on LLaMA and GPT-3.5/4 report that the systematic application of prompt principles yielded an average 57% increase in model output accuracy.

Education Next's (2025) linguistic clarity analysis revealed that a top-ranked mathematics program could not fully meet the 7 linguistic clarity criteria. This finding converges conceptually with PIF's semantic noise hypothesis and indicates that prompt ambiguity may be a structural problem in both human–human and human–machine communication. For the link to Bloom's (1968) Mastery Learning, see Part I, Table 3.

8. The Problem of Semantic Noise in Prompt Design

8.1. The Structural Ambiguity of Natural Language and Linguistic Multi-Interpretability

Natural languages ontologically conflict with the principle of perceptual invariance. In this study, linguistic multi-interpretability is used to describe the phenomenon whereby the same concept is encoded and decoded in numerous different ways across different languages, contexts, and individual minds; this phenomenon bears conceptual kinship with debates on linguistic relativity in linguistics. This view has also been treated, with different concepts, by Pinker (1994) and Crystal (2019): languages are shaped evolutionarily for the flexibility of everyday communication, not for scientific precision. In the case of LLMs, this inherited flexibility turns into a weakness. Bsharat, Myrzakhan, and Shen (2023, arXiv preprint) identified 26 prompt principles on the basis of experiments on LLaMA and GPT-3.5/4 models, reporting that the application of these principles yielded an average 57% accuracy increase under the experimental conditions of that study.

8.2. Hallucination: A Contemporary Semantic Pathology

Hallucination is defined as the production by LLMs of information that does not actually exist, with persuasive certainty (Lakera, 2026). Prompt ambiguity is not the sole cause of hallucination; model architecture, contradictions and gaps in training data, the stochastic nature of the token probability distribution (e.g., temperature and top-p settings), and inferences lacking a grounding base are also influential. Prompt ambiguity is only one of these factors; while the PIF-Prompt Model aims to reduce this factor, it cannot directly control in-model causes, and this should be recorded as a limitation. Among the main triggers at the prompt level are lack of context, contradictory instructions, and the absence of a verification mechanism. A strategy for reducing hallucination in educational settings should not be limited to authorizing the model to "say I don't know"; it should also be supported by connecting the model to an auditable, externally sourced knowledge pool (grounding). The PIF-Prompt Model can be positioned as a pedagogical component of this broader context-engineering framework. Among the concrete measures applicable at the prompt level are the following commands, which connect the model to a source and constrain fabrication: "Use only the texts/sources I have provided to you," "do not fabricate information you are unsure of; say 'I don't know' when necessary," and "use neutral, non-discriminatory language." Such commands can be added to the ethics/safety component of the PIF-Prompt Model's Clarity-Indexed Verification Layer.

In terms of the evidence hierarchy, peer-reviewed sources (Anowar and Khatun, 2026; Lee and Palmer, 2025) are used as primary evidence. Industry/practitioner reports (Datadog, 2025) are treated only as complementary observations and are not presented as direct evidence for the framework.

9. The PIF-Prompt Model: A Three-Layer Framework

In this section, PIF's three-pillar architecture is transformed into a three-layer operational model adapted to prompt engineering. The model's core proposition is the following: an educational prompt should consist of a Generalization Layer that builds context like an umbrella, a Uniqueness Layer that decomposes the output with barcode precision, and a Clarity-Indexed Verification Layer that audits the final product against an invariant reference set.

9.1. The Generalization Layer: The Umbrella Structure of Context

The Generalization Layer narrows the model's interpretive space and builds a common cognitive ground. Its function is the same as PIF's umbrella definition, summarized in the biology example as "Everything from a bacterium to an ostrich egg is a cell." Applied to a prompt, this layer should contain four basic components:

- **Role assignment:** The model's expertise frame is defined (e.g., "You are a 9th-grade biology teacher"). Bsharat et al. (2023) report that role assignment increases output consistency by an average of 18%.
- **Target audience definition:** The age, prior-knowledge level, and language proficiency of the student group the output is intended for are specified.
- **Conceptual framework:** The "common denominator" definition of the topic is given; subtypes or differences are not specified at this stage.
- **Pedagogical goal:** The learning objective the output will serve is written explicitly.

This structure converges with Sweller's (2024) strategy of reducing element interactivity. UNESCO's (2024a) AI competency framework for teachers emphasizes a similar requirement under the dimension of a "human-centered mindset."

9.2. The Uniqueness Layer: Barcoding the Output

The Uniqueness Layer defines the distinctive features of the expected output with fingerprint precision. As in PIF's "pencil versus fountain pen" example, each feature must be coded so as to prevent the model from confusing it. This layer consists of three main components.

Specifying the output format: whether it will be a bulleted list, a table, a flowchart, or plain paragraph must be defined precisely. **Quantitative constraints:** specifying precise limits such as "an explanation not exceeding 180 words" instead of "a long explanation," and "three different examples" instead of "a few examples," parallels PIF's principle of size invariance (Goldstone, 1998). **Linguistic and stylistic calibration:** the tone of the output, the terminological depth, and the criteria for selecting examples must be given explicitly.

9.3. The Clarity-Indexed Verification Layer

One of PIF's most innovative elements, the Clarity-Indexed Scoring System, posits that the value of an educational material is proportional not only to its difficulty but also to its clarity. This approach can be carried over to prompt engineering as a verification layer.

Score = Base × [(0.7 × p) + (0.3 × r)]

- **p (Clarity Weight) = 0.7** — the perceived clarity rate in the target audience
- **r (Discrimination Weight) = 0.3** — the purpose-directed level of discrimination

The weights 0.7 and 0.3 in the formula rest on a theoretical rationale; however, this uncertainty is explicitly recorded as a 'limitation.' These parameters must be recalibrated through inter-rater reliability studies.

In the first step, the model output is presented to a control group analogous to PIF's "White Paper Test." In the second step, the formula is applied; a low clarity score requires the prompt to be rewritten.

---

10. The EduCode Approach and Structured Prompt Templates

The EduCode Protocol proposed by PIF (see Part I, §4.2) calls for educational concepts to be represented by universal codes independent of natural language. This protocol is directly applicable to prompt design.

10.1. Comparison of PIF Principles with the Prompt-Engineering Literature

The table below summarizes the theoretical overlaps between PIF's three pillars, the 26 prompt principles of Bsharat et al. (2023), and UNESCO's (2024a, 2024b) competency frameworks.

Table 6 Comparison of PIF's Three Pillars with the Prompt-Engineering Literature

PIF Pillar	PIF Principle	Corresponding Principles in Bsharat et al. (2023)	UNESCO (2024a/b) Competency Dimension	Pedagogical Function
Generalization Layer	Present the most inclusive, indisputable 'umbrella' definition of the concept first; move to subtypes later.	Principle 1: Role assignment; Principle 6: Task definition; Principle 12: Specifying the target audience	Teacher: Human-centered mindset; Student: AI foundations	Reducing intrinsic load by establishing a cognitive anchor
Uniqueness Layer	Define the distinctive features of the expected output with fingerprint precision; specify format, tone, and constraints.	Principle 2: Format specification; Principle 14: Output length; Principle 16: Adding constraints; Principle 23: Providing examples (few-shot)	Teacher: AI applications; Student: AI techniques	Limiting output variance and hallucination
Clarity-Indexed Verification	Test the output in the target audience; Score = Base × [(0.7 × p) + (0.3 × r)]; target a clarity rate ≥ 95%.	Principle 22: Iterative refinement; Principle 26: Adding a verification step	Teacher: AI ethics and assessment; Student: Ethics and critical thinking	Ensuring the pedagogical validity of the output
EduCode Protocol	Eliminate natural-language ambiguity with a structured JSON/CNL template.	Principles 9, 10: Using a structured format	Teacher: Designing AI systems; Student: Digital literacy	Clearing semantic noise at the code level

Note. The table was compiled by the author by synthesizing Demirkuş (2026), Bsharat et al. (2023), and UNESCO (2024a, 2024b).

While Table 6 maps PIF's pillars to the relevant literature, Table 6b below offers a complementary, practitioner-oriented view, relating common prompt-writing techniques directly to the PIF-Prompt layers. In particular, techniques such as negative prompting (prohibitions) and chain-of-thought are positioned as functional components of the corresponding layers.

Table 6b Mapping Prompt-Writing Techniques to the PIF-Prompt Layers

Prompt-Writing Technique	**Function**	**Associated PIF-Prompt Layer**
Role assignment	Fixes the model's domain of expertise and point of view.	Generalization
Target-audience definition	Narrows the interpretive space by specifying age, prior knowledge, and language level.	Generalization
Context / conceptual frame	Provides the inclusive "common denominator" definition of the topic.	Generalization
Output format	Determines the structure—list, table, flowchart, etc.	Uniqueness
Length / quantity limit	Sets precise bounds such as "at most 180 words" or "three examples."	Uniqueness
Negative prompting (prohibitions)	Excludes undesired content (e.g., rote definitions, jargon, personal data).	Uniqueness
Verification / success criterion	Defines the clarity criterion against which the output will be tested.	Clarity-Indexed Verification
Chain-of-thought (CoT)	Guides the model to reason step by step and provide justification.	Clarity-Indexed Verification

Note. The table was created by the author to summarize the functional mapping between the three layers defined in §9 and common prompt-writing techniques; for chain-of-thought, see Wei et al. (2022) and Meincke et al. (2025).

10.2. JSON and Natural-Language Templates

Below are an advanced JSON template that follows EduCode logic and contains PIF's three layers, and a beginner-level natural-language template. The same structured-contract logic has been observed to be arrived at independently by users without formal JSON training in classroom practice; this serves as complementary evidence of the template's accessibility and intuitive learnability.

```
{  "role": "9th-grade biology teacher",

  "target_audience": {

    "age_group": "14-15",

    "prior_knowledge_level": "basic cell concept acquired",

    "language_proficiency": "Turkish B1"  },

  "general_framework": "Photosynthesis; the process by which

                  chlorophyll-bearing organisms convert light

                  energy into chemical energy",

  "pedagogical_goal": "Students can derive the input-output equation

                  of photosynthesis through reasoning",

  "uniqueness": {

    "format": "three-step flowchart",

    "word_limit": 220,
```

"number_of_examples": 2,

"tone": "explanatory, student-friendly",

"desired_output": ["current everyday language", "functional/conceptual definition"],

"undesired_output": ["Latin terminology only", "context-free rote definition"] },

"clarity_criterion": "when read to 5 students, at least 4 can write the equation correctly"}

The same content can also be expressed with natural English headings for users unfamiliar with the JSON format:

- **ROLE:** "You are a 9th-grade biology teacher."
- **TARGET AUDIENCE:** "Students aged 14–15 who have acquired the basic cell concept and are at the Turkish B1 level."
- **GENERAL FRAMEWORK:** "Photosynthesis; the process by which chlorophyll-bearing organisms convert light energy into chemical energy."
- **PEDAGOGICAL GOAL:** "Students can derive the input–output equation of photosynthesis through reasoning."
- **UNIQUENESS:** "A three-step flowchart, at most 220 words, 2 examples; do not use Latin terminology or rote definitions."
- **CLARITY CRITERION:** "When read to 5 students, at least 4 can write the equation correctly."

Meincke et al. (2025, arXiv preprint) report that the effect of the chain-of-thought principle diminishes in newer-generation models, whereas the effect of structured formats remains stable across all model generations. This finding provides complementary support that the PIF-Prompt Model carries model-generation-independent validity; it must be noted that the source in question is a non-peer-reviewed preprint and that validating this inference requires peer-reviewed studies.

10.3. Weak versus Strong Prompt Comparison: An Applied Example

To concretize the theoretical layers of the PIF-Prompt Model, a weak and a strong prompt written for the same learning objective are compared below. The table exemplifies the possible effect on output quality of whether or not the three layers (Generalization, Uniqueness, Clarity-Indexed Verification) are applied in a prompt. The "possible output quality" row reflects a theoretical expectation and must be tested empirically.

Table 7 Weak versus Strong Prompt Comparison (Photosynthesis Example)

Dimension	**Weak Prompt**	**Strong Prompt (PIF-Prompt Model)**
Prompt text	"Explain photosynthesis."	"As a 9th-grade biology teacher, prepare a three-step flowchart for students aged 14–15 who have acquired the basic cell concept, with which they can derive the input–output equation of photosynthesis through reasoning."
Generalization layer (context)	No role, target audience, or common-denominator definition; the interpretive space is broad.	Role, age/prior-knowledge level, and an inclusive conceptual frame are explicitly defined.
Uniqueness layer (constraints)	Format, length, and number of examples are unspecified; no prohibitions.	Flowchart format, at most 220 words, 2 examples; Latin terminology and rote definitions are avoided, and the output is verified against the expected core definition.
Verification layer (clarity)	Output is not tested; quality is not audited.	Tested with the clarity criterion "at least 4 of 5 students can write the equation correctly."
Possible output quality	Ambiguous, variable from person to person; high risk of hallucination and task drift.	Stable, low-variance, aligned with the pedagogical goal; auditable.

Note. The table was created by the author to exemplify the application of the PIF-Prompt Model's three layers. The difference in output quality in the last row is a theoretical expectation and requires validation through pilot applications.

The photosynthesis example in Table 7 can be extended to other subject areas. Table 7b below summarizes how the same weak→strong transformation logic operates in different contexts such as mathematics, biology assessment, and guidance counseling, exemplifying that the PIF-Prompt Model is not merely theoretical but a cross-disciplinary, applicable educational-engineering tool. Whereas in weak prompts the role, target audience, format, scope, and ethical limits are unspecified, in strong prompts these elements are structured through the Generalization, Uniqueness, and Clarity-Indexed Verification layers.

Table 7b Weak→Strong Prompt Transformations across Different Domains under PIF Principles (Summary)

Domain	Weak Prompt	Strong Prompt under the PIF-Prompt Model (summary)
Mathematics (derivative)	"Explain the derivative."	"As a 12th-grade mathematics teacher, explain the derivative to students who have acquired the concept of limits as the 'instantaneous rate of change,' using one graphical example and at most 200 words; verify the output against the criterion 'Can the student define the derivative in their own words?'"
Biology (assessment)	"Make a quiz on the cell."	"As a 9th-grade biology teacher, prepare 5 multiple-choice questions on cell organelles, each with a single correct answer, together with their expected core answers; questions should probe function rather than rote memory; do not use any personal student data."
Guidance (text)	"Write a text for exam anxiety."	"As a school guidance counselor, write a supportive informational text of at most 150 words for high-school students experiencing exam anxiety, containing three concrete coping suggestions; do not use real student names, numbers, or diagnostic information; do not give medical advice."

Note. The examples were created by the author to illustrate the cross-disciplinary application of the PIF-Prompt Model; all strong prompts are kept anonymous so as not to contain personal data. Expectations regarding output quality are theoretical and must be tested through pilot applications.

11. Educational Applications and Pedagogical Implications

11.1. Teacher Training: The Teacher as a "Semantic Engineer"

UNESCO's (2024a) teacher competency framework redefines the teacher's role in the age of AI under five dimensions: human-centered mindset, AI ethics, AI foundations and applications, AI pedagogy, and AI-supported professional development. The "semantic engineer" identity is located at the pedagogical intersection of these five dimensions. The proposed competency components are: (a) being able to apply the three layers of the PIF-Prompt Model simultaneously to in-class communication and LLM interaction, (b) being able to produce local prompt templates by adapting the principles of Bsharat et al. (2023) to the Turkish pedagogical context, (c) being able to assess the Clarity Index of outputs through peer review, (d) being able to recognize usage patterns that cross ethical boundaries (Kohnke & Zou, 2025; Lan et al., 2025).

11.2. Student Skills: Critical Prompt Literacy

UNESCO's (2024b) AI competency framework for students defines 12 competencies under four basic dimensions. The core skill the PIF-Prompt Model will impart to students is to treat GenAI outputs not as objects of passive consumption but of active audit. The proposed teaching sequence contains three stages: in the first stage, students concretely experience the concept of semantic noise by analyzing the outputs produced by their own ambiguous prompts. In the second stage, the logic of the White Paper Test is applied in the classroom. In the third stage, students develop metacognitive regulation skills by working with EduCode-style structured templates (Anowar and Khatun, 2026).

11.3. Assessment and Integration into the Curriculum

The Clarity-Indexed Scoring System proposed by PIF can be directly adapted for assessing prompt-engineering competency. The proposed rubric should contain three dimensions: (a) the adequacy of context construction (Generalization layer), (b) the precision of the output specifiers (Uniqueness layer), (c) the presence and applicability of a verification mechanism (Clarity layer). Each dimension is scored on a 0–10 scale.

Table 8 PIF-Prompt Competency Assessment Rubric (Example)

Dimension	0 Points	5 Points	10 Points
Generalization	Concept ambiguous; no common-denominator definition	Definition partly clear but missing one or more essential components	A single, inclusive, and indisputable definition is present
Uniqueness	No distinctive feature specified	Some distinctive features given	Clear distinctive features that prevent confusion are present
Verification	Output not tested	A limited/small check performed	Systematically tested with the Clarity Index

Note. Each dimension is scored on a 0–10 scale; intermediate values (e.g., 2, 7) are judged by linear interpolation. The rubric is an example application of the three dimensions (Generalization, Uniqueness, Verification) translated into concrete indicators and requires empirical calibration.

In the Turkish context, PISA 2022 data (OECD, 2023) reveal that part of the difficulty in students' mathematics scores concentrates in "converting context into mathematics." **An important caveat:** PISA achievement levels are multivariate; structural variables such as socioeconomic status, curriculum design, and teacher quality are determinative. The PIF-Prompt Model can offer value only as one component of this multifactorial picture.

11.4. Applied Example Modules: Transferring the PIF–EduCode Framework to Course Materials

To exemplify that the PIF-Prompt Model and the EduCode Protocol are applicable beyond being an abstract proposal, three example course modules designed according to the framework have been developed. These modules cover two biology topics—cellular energy processes and plant tissue systems—and in each, PIF's three layers (Generalization → Uniqueness → Clarity-Indexed Verification) and the EduCode concept codes operate simultaneously.

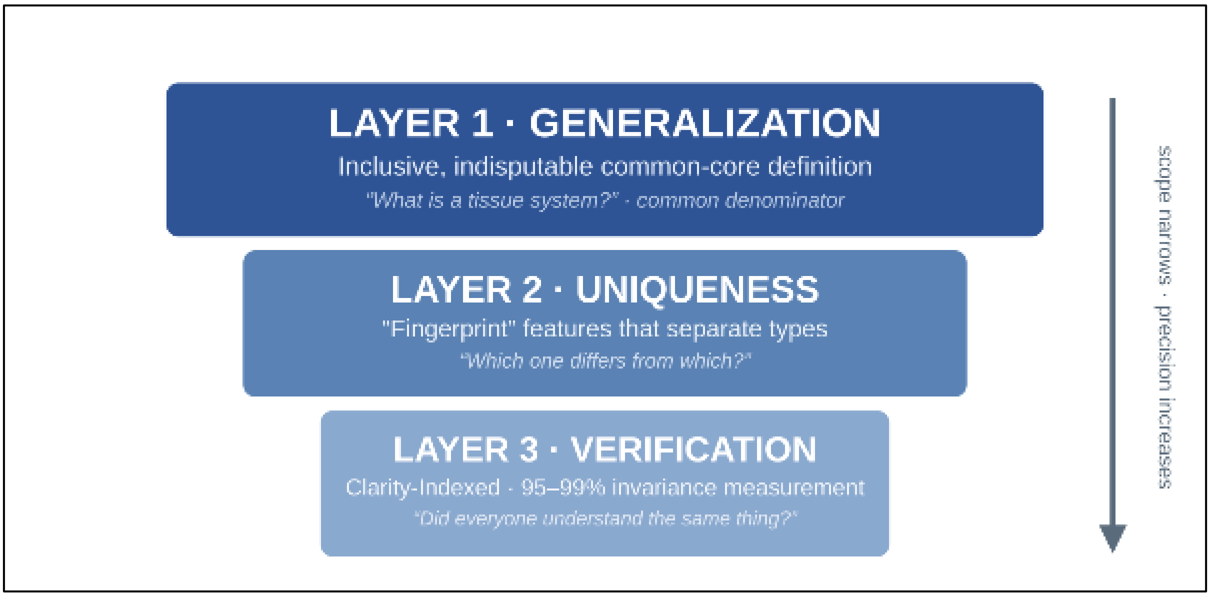

Figure 2 The three-layer architecture of the PIF-Prompt Model: as the scope narrows from top to bottom, semantic precision increases (Generalization → Uniqueness → Clarity-Indexed Verification).

The first module addresses photosynthesis and cellular respiration within a single cyclical framework (EduCode: BIO.007.2024 / BIO.008.2024). The Generalization layer unites both processes around the shared core of 'energy conversion'; the Uniqueness layer separates photosynthesis, aerobic respiration, and fermentation by their distinctive 'fingerprint' features; and the Verification layer measures the clarity rate with a short test of predefined expected core answers. By foregrounding the chemical equation ($6CO_2 + 6H_2O \rightarrow C_6H_{12}O_6 + 6O_2$) as a language-independent 'mathematical barcode,' the module constitutes a concrete classroom application of the EduCode record defined in §4.2.

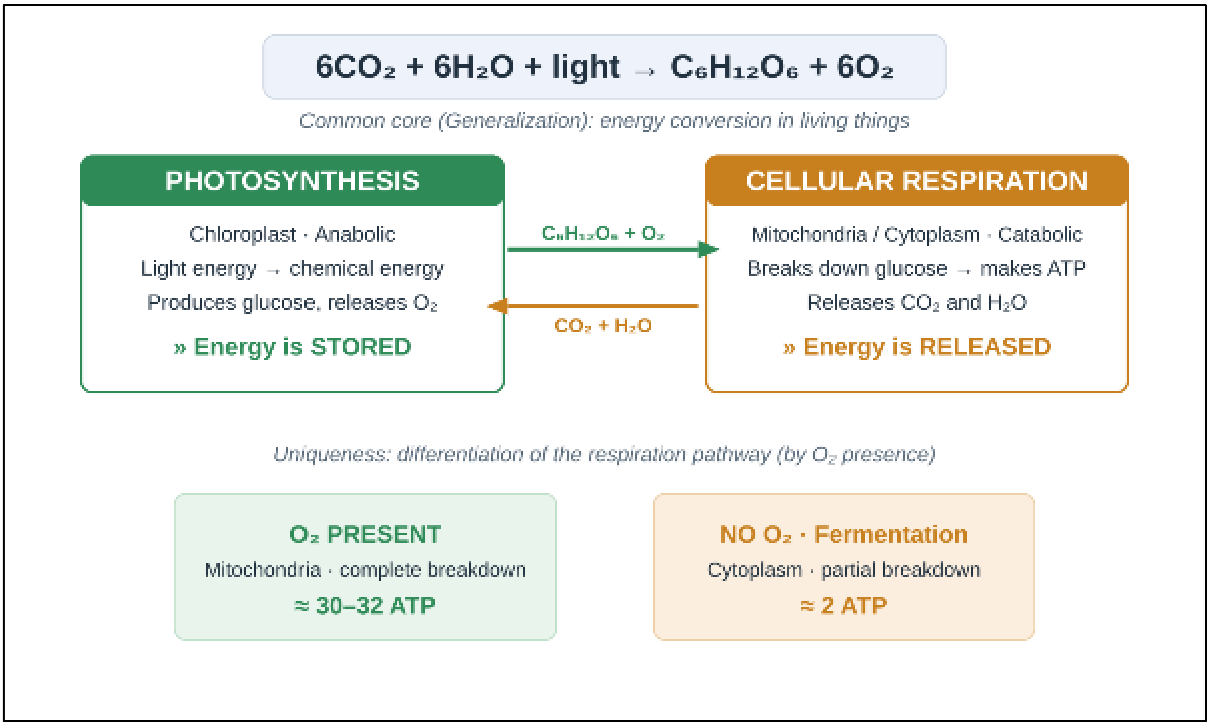

Figure 3 The photosynthesis–respiration module: energy conversion as the shared core (Generalization), the decomposition of the processes, and the unique distinction of the aerobic/anaerobic respiration pathway.

The distinctive aspect of this module is that it compares the photosynthesis and respiration processes within a single EduCode record and directly corrects a frequently encountered misconception (Table 9).

Table 9 EduCode Comparison Record of the Photosynthesis–Respiration Module

EduCode field	Content
Universal code (ID)	Photosynthesis: BIO.007.24 · Aerobic respiration: BIO.008.24.A · Fermentation: BIO.008.24.B
Mathematical language (invariant)	Aerobic respiration (A): $C_6H_{12}O_6 + 6O_2 \rightarrow 6CO_2 + 6H_2O + \sim 32$ ATP. Lactic acid fermentation (B): $C_6H_{12}O_6 \rightarrow$ 2 lactic acid + 2 ATP.
Frequently confused misconception and EduCode correction	Misconception: “Anaerobic and aerobic respiration occur in the same place.” Correction: aerobic respiration is completed in the mitochondria in eukaryotes; lactic acid and ethyl alcohol fermentation occur entirely in the cytoplasm.

Note. Adapted from the EduCode comparison record of the photosynthesis–respiration module (BIO.007–008).

The second and third modules address plant tissue systems under a shared parent code (BIO-BOT-DOKU-001) and define meristematic, ground, vascular, and dermal tissue with separate concept records. One of these is designed as an animated mini-test containing Clarity-Indexed short-answer questions and an accompanying mini-rubric; the other is a reusable template in which PIF layers and EduCode fields are explicitly mapped and which ends with a blank ‘copy-and-fill’ concept template. This template aims to facilitate the scalable application of the framework to different concepts.

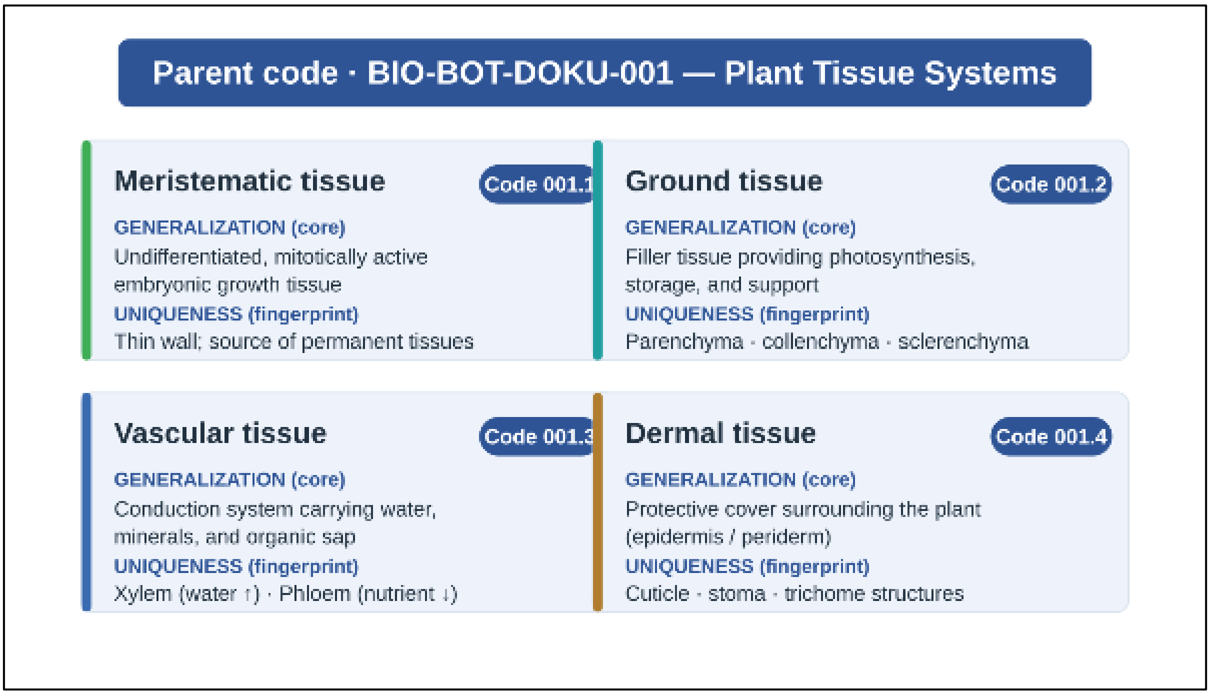

Figure 4 The plant tissue systems module: EduCode concept records under the parent code BIO-BOT-DOKU-001, with the generalization definition and unique distinctive features of the four tissues.

In the plant tissue modules, each tissue is represented not only with summary cards but also with a detailed comparison table containing cell state, wall structure, primary function, and typical location (Table 10).

Table 10 Detailed Comparison of Plant Tissue Systems

Tissue type · code	**Cell state**	**Wall and structure**	**Primary function**	**Typical location**
Meristem · 001.1	Undifferentiated, mitotically active	Thin primary wall; large nucleus, small vacuole	Longitudinal/lateral growth; production of new tissue	Apical meristem; lateral (cambium)
Parenchyma · 001.2	Living; some can re-divide	Thin primary wall; large vacuole, abundant space	Photosynthesis, storage, wound repair	Mesophyll, cortex, pith
Collenchyma · 001.2	Living; limited division	Irregularly thickened primary wall; non-lignified	Flexible mechanical support	Petiole, corners of young stems
Sclerenchyma · 001.2	Mostly dead	Thick, lignified secondary wall; fiber/stone cell	Rigid mechanical support	Around bundles, hard shell
Xylem · 001.3	Conducting elements dead (trachea/tracheid)	Thick lignified secondary wall; ring/spiral	Water + mineral conduction; support	Wood part of the bundles
Phloem · 001.3	Sieve element living, anucleate	Thin–medium primary wall; sieve plate	Conduction of organic matter (sugar)	Bast part of the bundles
Dermal · 001.4	Epidermis living; periderm mostly dead	Cuticle/stoma (epidermis); suberized cork (periderm)	Protection, reducing water loss, gas exchange	Surface of young organs; periderm in older stems

Note. Adapted from the detailed concept table of the plant tissue systems module (BIO-BOT-DOKU-001).

Three schematics that visualize the structural relationships of the tissues in these modules reinforce the uniqueness layer of the concepts: the placement of tissue in plant organs (Figure 5), the cell journey from meristem to permanent tissue (Figure 6), and source–sink transport in xylem and phloem (Figure 7).

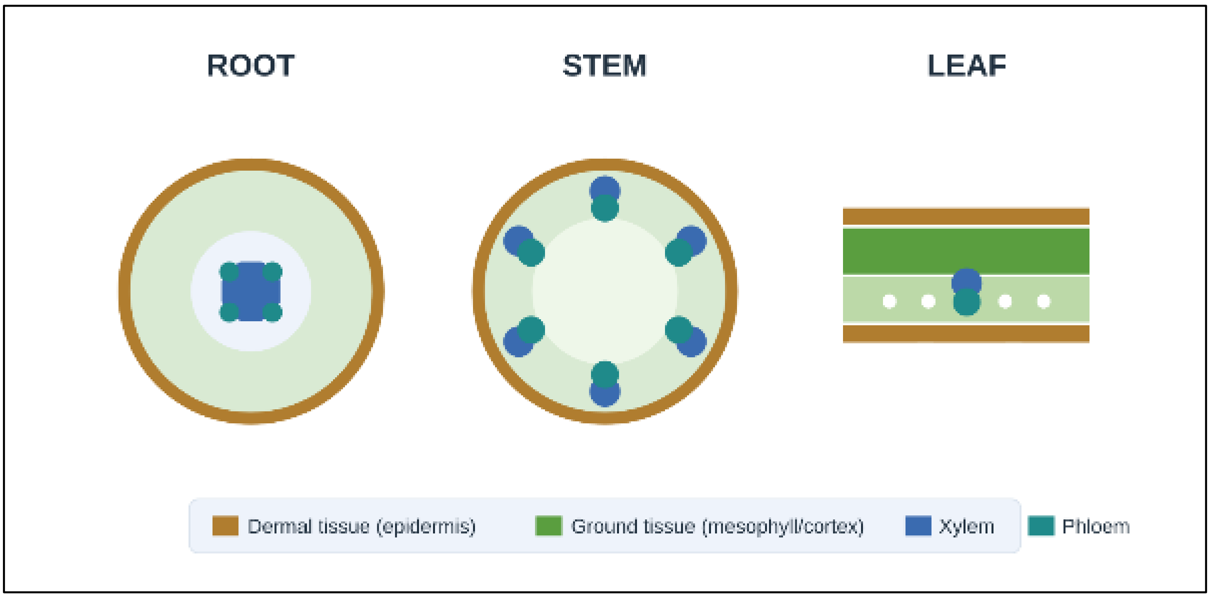

Figure 5 Schematic placement of the tissue systems in root, stem, and leaf cross-sections: dermal, ground, and vascular (xylem/phloem) tissues.

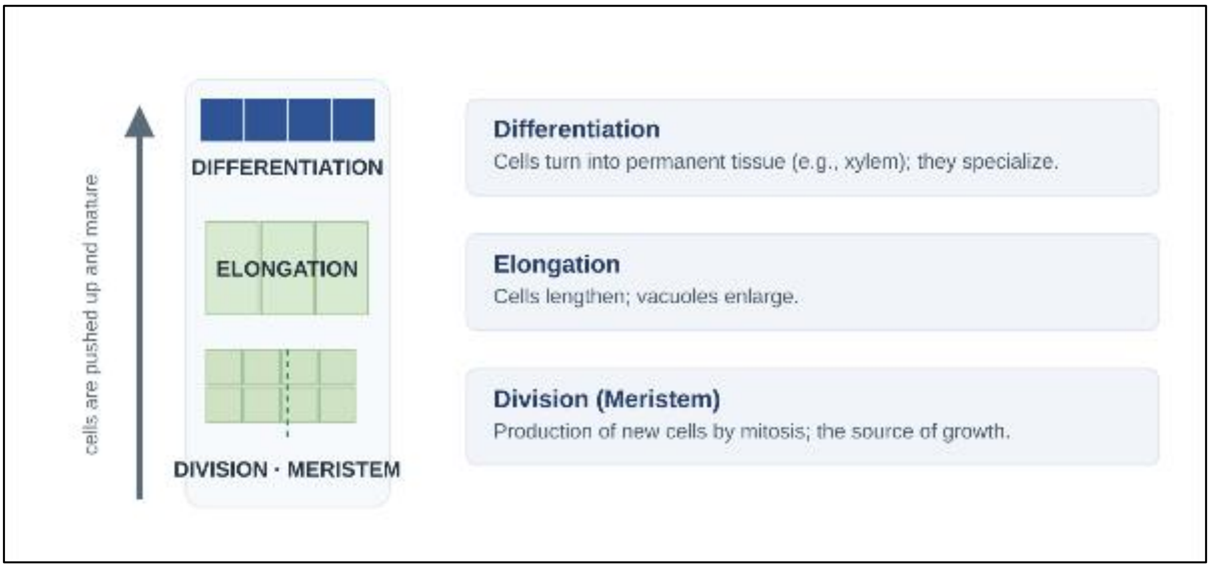

Figure 6 The cell journey from meristem to permanent tissue: division at the tip (meristem), then elongation and differentiation (static schematic of the animated module).

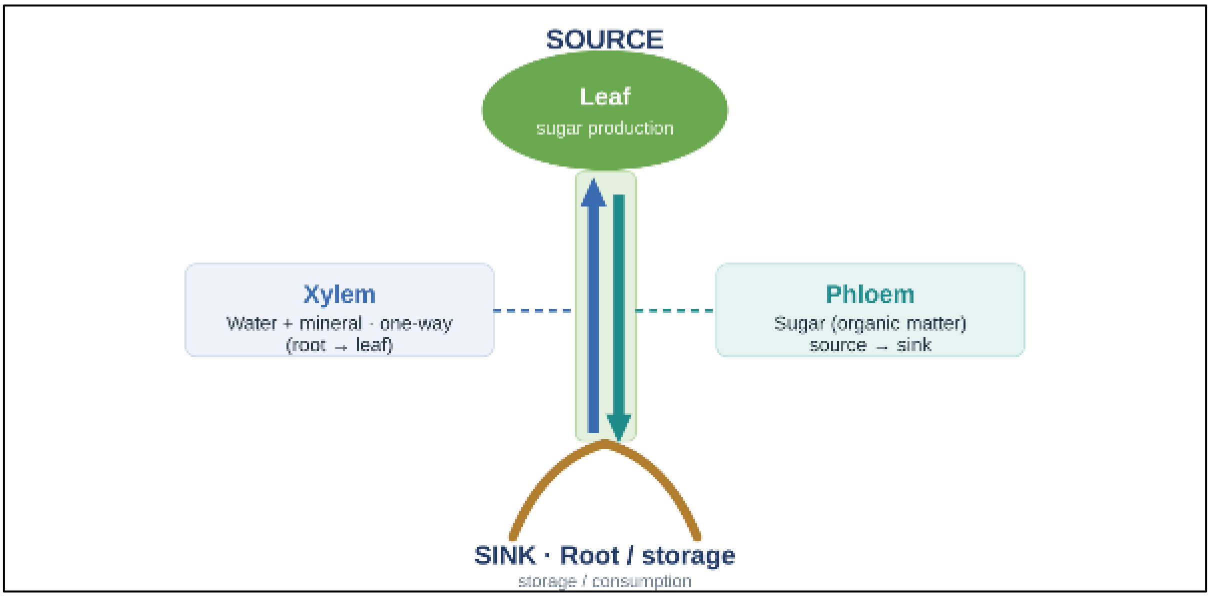

Figure 7 Source–sink transport in xylem and phloem: unidirectional water/mineral in the xylem, sugar conduction from source to sink in the phloem (direction varies by season).

All three modules share a common verification logic: concepts are presented to small groups with the "White Paper Test" approach, and for every concept code that falls below the targeted clarity threshold (≥ 95%), the relevant definition, table, or visual is reorganized. The PIF–EduCode correspondences summarized in Table 5 become operational in these modules. The structure of the modules is presented collectively in Table 11; the verification tools (rubric and example items) are presented in Appendix A.

Table 11 Structure of the Developed Example PIF–EduCode Course Modules

Module (EduCode code)	Concepts covered	Application of the PIF layers	Verification mechanism
Photosynthesis–Respiration interactive module (BIO.007.2024 / BIO.008.2024)	Photosynthesis, aerobic respiration, fermentation	Generalization: 'energy conversion' shared core; Uniqueness: the distinctive fingerprint of each process; Verification: mini-test with expected core answers	Short answer with the "White Paper Test" logic; any code below the ≥ 95% target is revised
Plant tissue systems — animated mini-test (BIO-BOT-DOKU-001)	Meristematic, ground, vascular, dermal tissue	Generalization definition and unique distinctive features for the four tissues; reinforcement with animated visuals	Clarity-Indexed short-answer questions and an accompanying mini-rubric
Plant tissue systems — reusable template (BIO-BOT-DOKU-001)	The same four tissues with a blank concept template	Explicit labeling of the three layers; mapping of PIF layers ↔ EduCode fields	A 'copy-and-fill' verification field with a 95–99% invariance target

Note. The modules are example/demonstration materials developed by the author according to the PIF-Prompt Model and the EduCode Protocol and made available in web-based form (nadidem.net). These materials exemplify the applicability of the framework and do not constitute empirical validation; their effectiveness must be tested through pilot studies in line with the limitations stated in §13.

12. Discussion: The Design of Mental Representation, with Ethical Considerations

This integrated analysis reveals that the Perceptual Invariance Framework is not merely a new exam-preparation technique, a material-development method, or a prompt-writing procedure; it constitutes a holistic approach to the design of mental representation that encompasses both human–human and human–machine educational communication.

12.1. Summary of Limitations and Core Research Questions

Before turning to the discussion, a brief summary of the study's main limitations is provided (for a detailed treatment, see §13). This study is the proposal of a theoretical framework; PIF, EduCode, and the PIF-Prompt Model—together with their numerical criteria—are not proven results but hypotheses to be tested. Concrete research questions stand out along three axes:

- **Psychometric validation:** What is the inter-rater reliability of Clarity-Indexed Scoring, and to what extent does the system converge with external learning criteria?
- **Cultural and linguistic validity:** Does mathematical/symbolic coding produce an equivalent semantic correspondence across different languages and cultures?
- **Feasibility and cost:** With what cost–benefit balance can the EduCode infrastructure scale in low-resource contexts without increasing the risk of digital inequality?

12.2. Key Findings and Strategic Recommendations

The integrated effect of the framework can be summarized in the inverse relationship between semantic noise and clarity: as one progresses from traditional material through the PIF filtering and coding stages, and from there to EduCode and PIF-Prompt output, a decrease in semantic noise is expected together with a rise in the level of clarity (Figure 8).

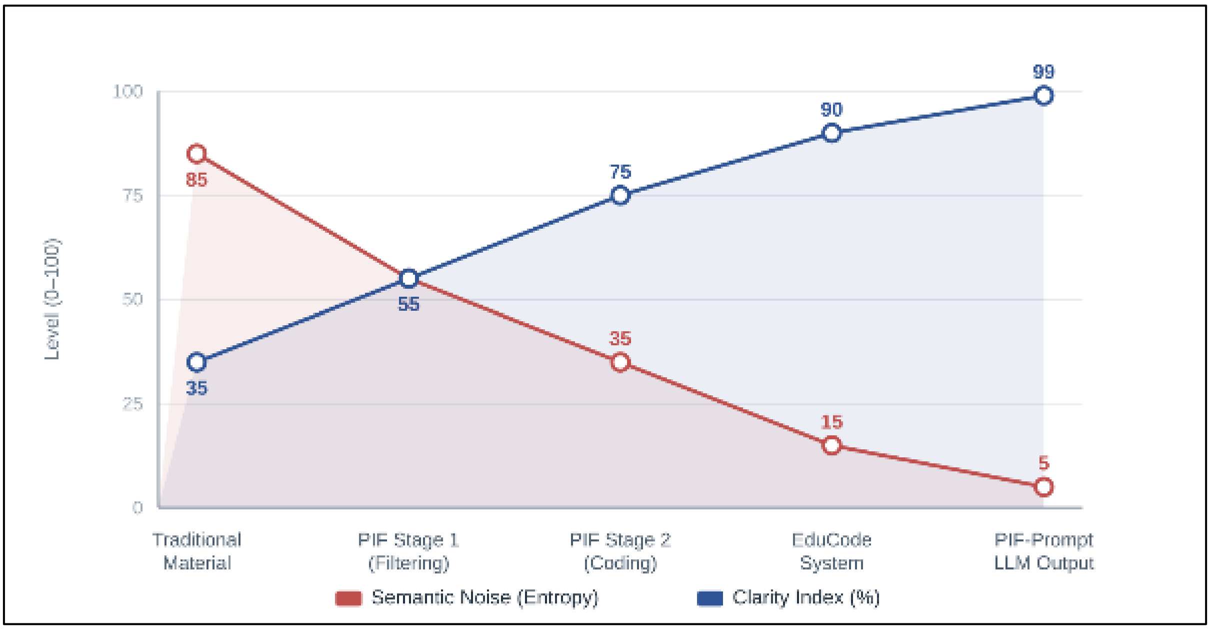

Figure 8 The theoretical inverse relationship between semantic noise and clarity along the PIF workflow. The values are not empirical measurements but conceptual/illustrative levels showing the proposed pattern.

- **Education Requires an Engineering Dimension That Complements Artistic Intuition:** The preparation of educational materials and prompts should not be left solely to the creator's subjective inspiration; the proposed framework suggests rethinking education as a high-standard, more engineering-oriented discipline that complements—rather than rejects—artistic intuition. The Generalization and Uniqueness rules must be calculated with the structural precision of this engineering dimension.
- **Justice Through Clarity:** An ambiguous item or prompt represents a systemic error; the cost of this error must be borne by the system, not the student. This requires a shift in emphasis from CTT's 'difficulty' paradigm to PIF's 'clarity' paradigm. There is no wholesale rejection of CTT here: CTT is a tradition optimized for selection- and ranking-oriented contexts (e.g., university entrance examinations) and is still necessary; PIF, in turn, offers a complementary framework for formative assessment and instructional material development. The Clarity Index should be positioned as a tool that builds an additional dimension on top of traditional item analysis rather than abolishing it.
- **Limitations of Natural Language:** To reduce semantic noise, core concepts can be transformed into 'Mathematical Barcodes' so that they can be retrieved at the level of cognitive automaticity.

- **Gradual Liberalization Model:** The system should implement 'Adaptive Invariance Levels' in which core conceptual foundations (Level 1) require 99% invariance, while higher-order applications (Level 3) allow 70% interpretive flexibility.

12.3. An Ethical Framework for Educational Engineering

- **Autonomy–Standardization Balance:** While invariant foundations ensure equity, pedagogical space for student choice, exploration, and personal meaning-making must be preserved.
- **Preservation of Cultural Diversity:** The EduCode system must incorporate cultural contextualization mechanisms. While mathematical representations remain invariant, examples should be culturally adaptable.
- **Inclusive Design for Special Needs:** The 99% invariance target must be interpreted with neurodiversity in mind. Materials should contain multiple access points (visual, auditory, kinesthetic).
- **Teacher Agency:** The engineering approach should empower rather than replace teacher expertise. Teachers should be trained as 'semantic engineers.'
- **Personal-Data Protection and Anonymization:** Writing identifying data such as real student names, school numbers, class information, or psychological status directly into prompts poses a risk under data-protection law (e.g., GDPR/KVKK), and such data may leak into the model's output or logs. Therefore the PIF-Prompt Model should be supported with negative prompts containing prohibitions such as "do not request personal data, do not carry personal data into the output, do not use real student names"; where needed, data should be anonymized (e.g., "a 9th-grade student") before use.
- **Centralization and Epistemic Pluralism:** Fixing meaning with universal codes and distributing it from a central repository (EduOntology; the "Central Bank of meaning") carries the risk of epistemic monopolization and ideological hegemony, especially in fields such as history, literature, social sciences, and philosophy where interpretation is inherently plural. The question of who decides, through what filter, the "common denominator" definition of a concept is inevitably political. To limit this risk, strict (mathematical/invariant) coding should be confined to core science/STEM concepts; in interpretive fields, EduCode should mark only the inclusive core and leave the interpretive layer explicitly plural and open to debate. Governance should be built on an open-access, version-controlled, and cross-institutionally pluralistic model rather than a single central authority (cf. §4.2).

---

13. Current Limitations

What Does This Study Not Prove? This article does not empirically prove the effect of PIF, EduCode, or the PIF-Prompt Model on learning outcomes. The clarity targets presented (≥ 95% applicable target, 99% theoretical ideal), the scoring weights (0.7 / 0.3), and the expected effect sizes are not proven values but initial assumptions that must be tested, especially for core invariant concepts. The study's contribution is limited to offering a testable theoretical framework, an applicable model, and concrete hypotheses; the validity of these propositions must be tested through the proposed multi-site pilot and controlled experimental studies.

This study is explicitly a proposal for a theoretical framework; none of the numerical values presented (e.g., the 99% Clarity Target, the 0.7 and 0.3 weights in the formula) is an experimentally proven universal parameter, but rather an initial proposal for operationalizing the framework. A limitation regarding the source hierarchy must also be recorded: policy magazines, industry reports, and preprints (e.g., Education Next, Lakera, etc.) are used in this study not as primary academic evidence but as complementary/grey literature; strengthening the core theoretical claims with peer-reviewed sources is left to future work.

- **Empirical Validation Gap:** The claim of reaching 99% comprehension fidelity and the effectiveness of the PIF-Prompt Model remain hypotheses requiring validation through large-scale, multi-site RCTs (N must be determined on the basis of a power analysis; Cohen's d must be reported).
- **Lack of Weight Justification:** No empirical basis is given for the weights w_1=0.7 and w_2=0.3 in the formula. This uncertainty is explicitly recorded as a limitation.
- **Assumption of Cultural and Linguistic Universality:** The assumption that mathematical/symbolic coding will exhibit cross-cultural equivalence is an open empirical question requiring international validation studies. The model was developed in the Turkish context, and its validity in international linguistic contexts has not been verified.
- **Implementation Complexity and Scalability:** This theory requires the complete redesign of curriculum materials (millions of pages), the updating of teacher-training programs (millions of educators globally), the transformation of assessment infrastructure (all standardized tests), and the restructuring of digital platforms

(EduCode integration across education systems). This undertaking is itself a multi-billion-dollar, decade-long process envisaged to be carried out without a comprehensive feasibility analysis or pilot implementation study.

- **Lack of Psychometric Validation:** The proposed Clarity-Indexed Scoring System represents a radical break from established psychometric paradigms and lacks validation studies. Several critical questions await answers: What is the inter-rater reliability of Clarity Index calculations? To what extent does the system converge with external learning criteria? What are the possible vulnerabilities to abuse or manipulation?
- **Teacher Resistance and Professional Identity:** Teachers trained in traditional constructivist or student-centered pedagogies may perceive the 'engineering' approach as mechanistic or reductionist. A deeper risk is the reduction of the teacher's intuitive, empathetic, and artistic classroom role to that of a data analyst who fills in JSON templates and computes clarity indices—leading, over the long term, to professional alienation and burnout and to the teacher becoming an interface of the EduCode platform. The 'semantic engineer' identity must therefore be positioned as a layer that strengthens, rather than alters, the teacher's situational intuition and professional autonomy.
- **Semantic Sterility and Cognitive Flexibility:** Completely eliminating ambiguity may also remove the productive difficulty that can serve as a catalyst in learning and scientific discovery. Standardizing concepts to 99% 'fingerprint' precision may reduce the learning process to a high-efficiency data transfer and weaken the cognitive flexibility and critical skepticism nourished by stretching concepts, generating new metaphors, and lateral thinking. In light of Bjork and Bjork's (2011) 'desirable difficulties' framework (see §7.1), PIF's noise elimination must observe a balance that preserves the germane difficulty that nourishes learning, and this balance must be tested empirically.
- **Technology Dependence and Digital Equity:** EduCode's dependence on digital infrastructure (QR codes, online databases, AI-supported ontologies) raises possible equity concerns in low-resource contexts, rural areas, and developing countries where internet connectivity and device access remain limited. The following alternatives are therefore proposed for low-technology conditions: a portable offline EduCode library (local database) requiring no internet; printed EduCode records in textbook supplements and printed concept cards; local versioned copies maintained at the school or district level. Thus QR codes can be designed to provide access to the same record from an offline source when there is no online connection.
- **Technical Scope:** The model is designed for educators and students; interventions at the level of LLM architecture are outside its scope.

14. Implementation Roadmap and Cost–Benefit Framework (2025–2035)

- **Stage 1 (2025–2027):** Development of comprehensive EduCode ontologies for STEM; training of 1,000 pilot teachers as 'Educational Engineers'; conducting multi-site RCTs with more than 10,000 students.
- **Stage 2 (2028–2031):** Integration of EduCode into the national curricula of 5 pioneer countries; development of adaptive learning platforms containing Perceptual Invariance algorithms.
- **Stage 3 (2032–2035):** Implementation of PIF standards in more than 30 countries; development of AI-supported real-time semantic noise detection in classrooms; creation of a global EduCode repository hosting more than 1 million invariant concepts.

14.1. Cost–Benefit Analysis Framework

- **Initial Investment (2025–2030):** Research and development: USD 50–100 million; teacher training: USD 200–500 million; technology infrastructure: USD 100–300 million. Total: USD 350–900 million globally.
- **Expected Benefits:** Theoretically predicted (hypothetical scenario): a 20–40% improvement in international assessment scores (PISA, TIMSS); a 30–50% reduction in achievement gaps; a USD 2–5 return per USD 1 spent (rough/illustrative estimates based on the general education return-on-investment literature). These figures are order-of-magnitude illustrative estimates and require validation through detailed feasibility and cost modeling.

14.2. Future Research Directions

Directions proposed for future research: (a) testing adaptability across different disciplines, (b) adaptation for students with disabilities and neurodivergent students, (c) integration with the AI-ethics dimension, (d) policy work toward transforming EduCode into an open-access international standard.

14.3. Recommendations and Implementation Protocol

This section turns the Perceptual Invariance Framework (PIF) principles grounded theoretically in the earlier parts into directly applicable recommendations at the level of lecture notes, presentations, prompts, materials, sample questions and exercises, and then binds them to a sustainable implementation protocol at the classroom and institutional levels. Proceeding from the claim that semantic noise constitutes a shared pathology in both human and machine-based communication (Shannon & Weaver, 1949), the recommendations are synthesized with Cognitive Load Theory (Sweller, 2024; Sweller, van Merriënboer & Paas, 2019), Information/Message Design (Pettersson, 2002), the Cognitive Theory of Multimedia Learning (Mayer, 2009) and the formative-feedback literature (Hattie & Timperley, 2007; Shute, 2008; Wisniewski, Zierer & Hattie, 2020). All recommendations are not proven prescriptions but design principles to be tested through pilot studies within the study's overall limitations (Section 13).

14.3.1. Design-Level Recommendations

The following six recommendations adapt PIF's three layers (Generalization → Uniqueness → Clarity-Indexed Verification) to the core components of teaching practice. The shared aim is to reduce the extraneous/semantic noise arising from a coding error between intent and perception while preserving the germane load that arises from cognitively wrestling with a concept.

- **Lecture notes.** Each concept should be given first as a single, plain, most-inclusive umbrella definition (Generalization) and then through the fingerprint features distinguishing it from similar concepts (Uniqueness). The umbrella definition serves as a stable anchor onto which new information attaches; distinctive features are not buried in prose but presented side by side via bulleted lists or comparison tables. Contiguous comparison and systematic variation lighten cognitive load and sharpen conceptual boundaries, while terminological fidelity (EduCode/Controlled Natural Language) is preserved (Pettersson, 2002; Chen, Paas & Sweller, 2023; Kuhn, 2014).
- **Presentations.** Slides should be divided into three sequential layers separated by clear boundaries, so that the definition (Generalization), the difference (Uniqueness) and the assessment (Verification) are not crammed onto one slide. In the Generalization layer only a single central message appears on screen and decorative/non-functional visuals are removed (coherence principle). The Uniqueness layer presents distinctive features via a matrix/mind map. In the Verification layer a short "White Paper Test" is applied; if about 95% of the class does not arrive at the same response, the flaw lies in the slide and the concept is revised immediately (Mayer, 2009; Sweller, 2024; Shute, 2008).
- **Prompts.** Vague, context-free prompts produce hallucination and conceptual drift in large language models; prompts should be built at Controlled-Natural-Language precision in three structures: (a) Generalization—role/context and audience level; (b) Uniqueness—constraints and negative prompts specifying format, word limit, tone, concepts to avoid, and a ban on personal data; (c) Clarity/verification criterion—the output's acceptance criterion. This turns the interaction into a repeatable, low-variance engineering tool (Lee & Palmer, 2025; Qian, 2025; Kuhn, 2014).
- **Material preparation (LLM-as-a-judge).** Material should pass a semantic-noise test before reaching the learner. The teacher gives the material to an independent model with a rubric; the model is asked not to generate text but to judge it against predefined clarity, readability and consistency criteria, flagging polysemous/complex passages and suggesting simplifications. Because models may overestimate difficulty and struggle to hit a precise level, they should be used as a complement to—not a replacement for—human review (Paraschiv, Dascalu & Solnyshkina, 2025).
- **Sample questions.** Questions should be evaluated not only by difficulty and discrimination but also by clarity: a good item evokes the same stable response in the large majority of the audience. The stem is cleansed of natural-language entropy (interpretive isolation); distractors are built from the genuine fingerprints of other concepts; and in piloting or LLM-as-a-judge testing at least 95% of the audience should agree on what the question asks (Haladyna, Downing & Rodriguez, 2002; Shute, 2008).
- **Exercises.** Feedback should not remain at the "right/wrong" level but be delivered through a three-stage mechanism: (a) diagnostic—whether the error is at the Generalization or Uniqueness stage; (b) scaffolding + schema repair—re-anchoring the core definition or pointing to the distinctive difference; (c) mastery verification—an isomorphic new question. Feedback has a medium average effect (d=0.48) that depends largely on the information conveyed, and should be supportive, timely, specific and task-focused (Wisniewski, Zierer & Hattie, 2020; Hattie & Timperley, 2007; Shute, 2008).

Table 12 Mapping of Design Recommendations to PIF Layers and Sources

Recommendation area	PIF layer/principle	Concrete application criterion	Principal basis
Lecture notes	Generalization + Uniqueness	Dual definition; counter-example differentiation; terminological fidelity	Pettersson (2002); Chen, Paas & Sweller (2023); Kuhn (2014)
Presentations	Visual separation of the three layers	One focus per layer; matrix/mind map; closing verification	Mayer (2009); Sweller (2024); Shute (2008)
Prompts	Generalization + Uniqueness + Verification	Role/context; constraints and negative prompts; verification criterion	Lee & Palmer (2025); Qian (2025); Kuhn (2014)
Material preparation	Clarity-Indexed Verification	Clarity/readability audit via LLM-as-a-judge	Paraschiv, Dascalu & Solnyshkina (2025)
Sample questions	Clarity (p) + discrimination (r)	Interpretive isolation; fingerprint-based distractors; ≥95% clarity	Haladyna, Downing & Rodriguez (2002); Shute (2008)
Exercises	Verification + feedback	Diagnostic feedback; schema repair; isomorphic verification	Wisniewski, Zierer & Hattie (2020); Hattie & Timperley (2007)

Note. The table maps the design recommendations to PIF layers and supporting literature. Numerical criteria (e.g., ≥95% clarity) are not proven parameters but initial assumptions to be calibrated through pilot studies (cf. Section 13).

14.3.2. PIF Implementation Protocol: Operational and Institutional Recommendations

The following recommendations bind the design principles above to a sustainable, auditable and reusable protocol at the classroom and institutional levels, so that PIF does not remain an abstract framework but is transferred into daily teaching practice and institutional workflows.

- **Retrieval, spaced repetition and corrective feedback.** A durable conceptual response is not formed merely by reading; exercises should be structured with low-stakes short quizzes, spaced repetition and corrective feedback. Retrieval practice and spaced repetition markedly strengthen long-term retention (Roediger & Karpicke, 2006; Dunlosky et al., 2013).
- **A "PIF Implementation Checklist."** A common checklist should be developed for lecture notes, presentations, prompts, questions and exercises; each instructional unit should be systematically audited against PIF principles before release (see Table 13).
- **Adapting the EduCode template to all materials.** A simple EduCode record (code, name, universal name, generalization definition, uniqueness features, related concepts, misconception risk, sample question and expected core answer) should be created for every concept (cf. §11.4, Appendix A, Table 16). All materials are thus connected to the same conceptual database and the Controlled-Natural-Language logic (Kuhn, 2014) is carried across the ecosystem.
- **Sourcing and verification in AI-assisted materials.** Because generative AI carries the risk of hallucination and fabricated sources, every AI-assisted content should state which sources it relied on, which concept was used with which definition, and which expert review it passed; outputs should be reviewed by domain, assessment and language experts (UNESCO, 2024a). The verification layer (§9.3) is a mandatory ethical/safety layer here.
- **Educating students as meaning auditors.** Students should be asked not only to query AI but to evaluate the response against generalization, uniqueness and verification criteria. Context, clarity, exemplification and task constraints are known to improve LLM accuracy (Bsharat, Myrzakhan & Shen, 2023; Cain, 2023; Sahoo et al., 2024).
- **Empirical testing through pilot studies.** Materials prepared per PIF should be compared with traditional materials and tested for comprehension accuracy, response variance, cognitive load, delayed retention and satisfaction. This testing is directly tied to Section 13 and §14.2 and moves PIF toward a measurable model.

Table 13 PIF Implementation Checklist (Instructional-Unit Audit)

Audit question	Related PIF component
Is the concept's inclusive generalization (common-denominator) definition provided?	Generalization Rule
Is the uniqueness (fingerprint) definition distinguishing it from similar concepts provided?	Uniqueness Rule
Is the risk of confusion with similar concepts explicitly stated?	Uniqueness / Semantic Noise
Are context, task, output format and verification criterion explicit in the prompt?	PIF-Prompt Model
Is the expected core answer of the sample question defined?	Clarity-Indexed Scoring
Does the exercise include a clear, actionable feedback mechanism?	Formative Feedback

Note. The checklist was created by the author to audit instructional units for conformity with PIF principles before release.

In sum, these recommendations operationalize PIF's theoretical goal—perceptual invariance, the same concept understood the same way by different people—across the six core components of teaching practice and the institutional workflow. Lecture notes and presentations filter cognitive load via Controlled-Language and multimedia principles; LLM-as-a-judge scans for semantic noise that escapes the human eye; questions and exercises operationalize the Mastery Learning philosophy through formative feedback. The educator's role is not to reject creativity but to use it, under PIF's engineering frame, to design systems that minimize semantic noise. The effectiveness of these recommendations must likewise be tested empirically within the study's overall limitations (Section 13).

15. Conclusion

The Perceptual Invariance Framework does not ignore individual differences; on the contrary, it acknowledges them and manages them through “combination” methods in order to unite all learners at the “Common Denominator” of objective truth. This study has shown that PIF’s core conceptual tools can be systematically adapted both to educational material design (Part I) and to educational communication established with generative AI (Part II). The proposed PIF-Prompt Model offers the potential to move prompt engineering from trial-and-error processes toward a theoretically framed practice.

The model’s three layers—the Generalization Layer, in which context is built like an umbrella; the Uniqueness Layer, in which output is decomposed with barcode precision; and the Clarity-Indexed Verification Layer—combine pedagogically rooted principles with the empirical findings of the AI literature. Reversing the failures observed in PISA and World Bank data requires not an increase in the quantity of testing but a revolution in the quality of the “language of instruction.” Resonating with Bloom’s (1968) Mastery Learning tradition, and supported by Education Next’s (2025) empirical findings and Kuhn’s (2014) CNL tradition, this vision carries the potential to evolve from a local theoretical proposition into a universal educational standard.

In sum, this framework operates at three complementary levels: PIF represents the engineering of meaning; EduCode represents the coded educational application of this engineering of meaning; and the PIF-Prompt Model represents the adaptation of the same principle to human–AI communication.

Finally, it must be emphasized that this study proposes a theoretical model; all the components presented—the Perceptual Invariance Framework, the EduCode Protocol, and the PIF-Prompt Model, together with their associated numerical criteria—are not proven results but propositions to be tested. The validity of the model must be tested through multi-site pilot applications and controlled experimental studies.

Compliance with ethical standards

Acknowledgments

In the preparation of this article, generative AI-supported tools (Claude, ChatGPT, Gemini, Perplexity AI) were used under a principle of transparency for: (a) concept mapping and theoretical framework construction; (b) checking of writing conventions; (c) English–Turkish terminology consistency control. These tools were used only at an auxiliary level; all generated text was reviewed and verified by the author. The theoretical claims, analyses, and final academic responsibility of the study rest entirely with the author.

Disclosure of conflict of interest

The author declares no conflict of interest. The PIF-Prompt Competency Scale and the EduCode Protocol are shared open-access for academic use, and the author has no direct commercial interest in them; should commercial applications arise in the future, this declaration will be updated.

Statement of ethical approval

Because this study is a theoretical/model-developing work that did not involve human participants, clinical interventions, personal data, or identifiable information, ethical approval and informed consent were not required for the present manuscript. A separate ethics committee application (Van YYU Scientific Research Publication Ethics Committee for Social and Human Sciences) is under way for the development of the related PIF-Prompt Competency Scale, which constitutes an independent study. This study aligns with the United Nations Sustainable Development Goals, particularly SDG-4 (Quality Education).

Visual Materials and Accessibility

All figures in the article (Figures 1–8) were created by the author. The tables were adapted from the author's related works and from the sources indicated in the text, with the source noted beneath each table. The EduCode example modules and the materials related to the PIF-Prompt Competency Scale are provided open-access at nadidem.net.

References

[1] Anderson, J. R. (2020). Cognitive psychology and its implications (9th ed.). Worth Publishers.

[2] Anowar, S., & Khatun, R. (2026). Prompt engineering in education: A systematic review of theory, applications, and future directions. Journal of Education, Society & Sustainable Practice, 2, 30–42. https://doi.org/10.63697/jessp.2026.10093

[3] Ausubel, D. P. (1960). The use of advance organizers in the learning and retention of meaningful verbal material. Journal of Educational Psychology, 51(5), 267–272. https://doi.org/10.1037/h0046669

[4] Bartlett, F. C. (1932). Remembering: A study in experimental and social psychology. Cambridge University Press.

[5] Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough, & J. R. Pomerantz (Eds.), Psychology and the real world: Essays illustrating fundamental contributions to society (pp. 56–64). Worth Publishers.

[6] Bloom, B. S. (1968). Learning for mastery. Evaluation Comment, 1(2), 1–12.

[7] Bsharat, S. M., Myrzakhan, A., & Shen, Z. (2023). Principled instructions are all you need for questioning LLaMA-1/2, GPT-3.5/4 (arXiv:2312.16171). arXiv. https://doi.org/10.48550/arXiv.2312.16171

[8] Cain, W. (2023). Prompting change: Exploring prompt engineering in large language model AI and its potential to transform education. TechTrends, 68, 47–57. https://doi.org/10.1007/s11528-023-00896-0

[9] Chen, O., Paas, F., & Sweller, J. (2023). The relationship between interleaving and variability effects: A cognitive load theory perspective. Education Sciences, 13(11), 1138. https://doi.org/10.3390/educsci13111138

[10] Crystal, D. (2019). The Cambridge encyclopedia of the English language (3rd ed.). Cambridge University Press. https://doi.org/10.1017/9781108556965

[11] Datadog. (2025). Detecting hallucinations with LLM-as-a-judge: Prompt engineering and beyond. Datadog Engineering Blog. https://www.datadoghq.com/blog/ai/llm-hallucination-detection/ [Industry observation; not a peer-reviewed source.]

[12] Demirkuş, N. (2024a). Natural and universal mathematics in science teaching, learning and education — Lecture notes (unpublished lecture notes). https://nadidem.net/ders/bmat.html

[13] Demirkuş, N. (2024b). Special teaching methods II (unpublished lecture notes). https://nadidem.net/ders/K12.htm

[14] Demirkuş, N. (2026). Educational engineering in light of the perceptual invariance framework: Semantic noise elimination and universal mathematical language construction. World Journal of Advanced Engineering Technology and Sciences, 18(03), 081–097. https://doi.org/10.30574/wjaets.2026.18.3.0131

[15] Dewey, J. (1910). How we think. D.C. Heath.

[16] Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. Psychological Science in the Public Interest, 14(1), 4–58. https://doi.org/10.1177/1529100612453266

[17] Education Next. (2025). What AI revealed about a top math program. Education Next. https://www.educationnext.org/what-ai-revealed-about-a-top-math-program/

[18] Gibson, E. J. (1969). Principles of perceptual learning and development. Appleton-Century-Crofts.

[19] Goldstone, R. L. (1998). Perceptual learning. Annual Review of Psychology, 49, 585–612. https://doi.org/10.1146/annurev.psych.49.1.585

[20] Guskey, T. R. (2024, August). Empowering students through mastery learning. Corwin Connect. https://corwin-connect.com/2024/08/empowering-students-through-mastery-learning/

[21] Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. Applied Measurement in Education, 15(3), 309–334. https://doi.org/10.1207/S15324818AME1503_5

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

[23] Klein, F. (1872). Vergleichende Betrachtungen über neuere geometrische Forschungen (Erlangen Program). Andreas Deichert.

[24] Kohnke, L., & Zou, D. (2025). Artificial intelligence integration in TESOL teacher education: Promoting a critical lens guided by TPACK and SAMR. TESOL Quarterly. https://doi.org/10.1002/tesq.3396

[25] Kuhn, T. (2014). A survey and classification of controlled natural languages. Computational Linguistics, 40(1), 121–170. https://doi.org/10.1162/COLI_a_00168

[26] Laixi. (2025). LLM interview series (8): Prompt engineering — Principles, patterns, and best practices. Medium. https://medium.com/@huanzidage/llm-interview-series-8-prompt-engineering

[27] Lakera. (2026). The ultimate guide to prompt engineering in 2026. Lakera Blog. https://www.lakera.ai/blog/prompt-engineering-guide

[28] Lan, G., Feng, X., Du, S., Song, F., & Xiao, Q. (2025). Integrating ethical knowledge in generative AI education: Constructing the GenAI-TPACK framework for university teachers' professional development. Education and Information Technologies, 30(11), 15621–15644. https://doi.org/10.1007/s10639-025-13427-6

[29] Lee, D., & Palmer, E. (2025). Prompt engineering in higher education: A systematic review to help inform curricula. International Journal of Educational Technology in Higher Education, 22, Article 7. https://doi.org/10.1186/s41239-025-00503-7

[30] Leibniz, G. W. (1666). Dissertatio de arte combinatoria.

[31] Logan, G. D. (1988). Toward an instance theory of automatization. Psychological Review, 95(4), 492–527. https://doi.org/10.1037/0033-295X.95.4.492

[32] Mayer, R. E. (2009). Multimedia learning (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511811678

[33] Meincke, L., Mollick, E. R., Mollick, L., & Shapiro, D. (2025). Prompting science report 2: The decreasing value of chain of thought in prompting (arXiv:2506.07142). arXiv. https://doi.org/10.48550/arXiv.2506.07142

[34] OECD. (2023). PISA 2022 results (Vol. I): The state of learning and equity in education. OECD Publishing. https://doi.org/10.1787/53f23881-en

[35] Paraschiv, A., Dascalu, M., & Solnyshkina, M. (2025). Assessing the readability of Russian textbooks using large language models. Information, 16(12), Article 1071. https://doi.org/10.3390/info16121071

[36] Pettersson, R. (2002). Information design: An introduction. John Benjamins Publishing Company. https://doi.org/10.1075/ddcs.3

[37] Pinker, S. (1994). The language instinct. William Morrow.

[38] Qian, Y. (2025). Prompt engineering in education: A systematic review of approaches and educational applications. Journal of Educational Computing Research, 63(7–8), 1782–1818. https://doi.org/10.1177/07356331251365189

[39] Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. Trends in Cognitive Sciences, 20(9), 676–688. https://doi.org/10.1016/j.tics.2016.07.002

[40] Roediger, H. L., III, & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. Psychological Science, 17(3), 249–255. https://doi.org/10.1111/j.1467-9280.2006.01693.x

[41] Sahoo, P., Singh, A. K., Saha, S., Jain, V., Mondal, S., & Chadha, A. (2024). A systematic survey of prompt engineering in large language models: Techniques and applications (arXiv:2402.07927). arXiv. https://doi.org/10.48550/arXiv.2402.07927

[42] Schön, D. A. (1983). The reflective practitioner: How professionals think in action. Basic Books.

[43] Shannon, C. E., & Weaver, W. (1949). The mathematical theory of communication. University of Illinois Press.

[44] Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. Psychological Review, 84(2), 127–190. https://doi.org/10.1037/0033-295X.84.2.127

[45] Shute, V. J. (2008). Focus on formative feedback. Review of Educational Research, 78(1), 153–189. https://doi.org/10.3102/0034654307313795

[46] Stadler, M., Bannert, M., & Sailer, M. (2024). Cognitive ease at a cost: LLMs reduce mental effort but compromise depth in student scientific inquiry. Computers in Human Behavior, 160, Article 108386. https://doi.org/10.1016/j.chb.2024.108386

[47] Sweller, J. (2024). Cognitive load theory and the curriculum. In Research handbook on curriculum and education (pp. 250–264). Edward Elgar Publishing. https://doi.org/10.4337/9781802208542.00017

[48] Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. Educational Psychology Review, 31, 261–292. https://doi.org/10.1007/s10648-019-09465-5

[49] Turgut, M. F., & Baykul, Y. (2019). Eğitimde ölçme ve değerlendirme [Measurement and evaluation in education] (Extended ed.). Pegem Akademi.

[50] UNESCO. (2021). Global education monitoring report 2021/2: Non-state actors in education. UNESCO Publishing. https://doi.org/10.54676/UZQV8501

[51] UNESCO. (2024a). AI competency framework for teachers. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000391104

[52] UNESCO. (2024b). AI competency framework for students. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000391105

[53] Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E. H., Le, Q. V., & Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. Advances in Neural Information Processing Systems, 35, 24824–24837.

[54] Winget, M., & Persky, A. M. (2022). A practical review of mastery learning. American Journal of Pharmaceutical Education, 86(10), Article 8906. https://doi.org/10.5688/ajpe8906

[55] Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. Frontiers in Psychology, 10, Article 3087. https://doi.org/10.3389/fpsyg.2019.03087

[56] World Bank. (2022). The state of global learning poverty: 2022 update. World Bank Group. https://www.worldbank.org/en/topic/education/publication/state-of-global-learning-poverty

[57] World Bank. (2024). Türkiye poverty and equity brief. Washington, DC. https://documents.worldbank.org/en/publication/documents-reports/documentdetail/099850001032526747

Appendix A. Assessment Tools for the Example Modules

Below are the assessment tools that make operational the Clarity-Indexed Verification layer of the example modules introduced in §11.4. Table 14 presents the short-answer rubric that scores, at three levels, the degree to which student responses coincide with the core meaning of the concept; Table 15 presents the example verification items used together with this rubric and the predefined expected core answer for each item.

Table 14 Clarity-Indexed Short-Answer Assessment Rubric

Level	Score	Description
Full match	2	The answer correctly and completely contains the core meaning of the concept; the main distinctive feature(s) are correct; there is no scientific error; the expression is clear and unambiguous.
Partial match	1	The answer contains the basic aspect of the concept but an important point is missing or unclear; there are minor conceptual shifts or expression problems.
No match / erroneous	0	The answer has deviated from the concept, does not carry the core meaning, or contains a serious conceptual error.

Note. Adapted from the clarity-indexed short-answer rubric of the tissue module; for each concept, the targeted criterion is a class-level clarity rate of ≥ 95%.

Table 15 Example Verification Items and Expected Core Answers

Question (summary)	Target EduCode code	Expected core answer
What does "tissue system" mean? (with reference to the tissue and organ level)	BIO-BOT-DOKU-001	A tissue system is the organized arrangement, within plant organs, of several tissue types with similar or related functions; for example, the vascular tissue system is the whole formed by the xylem and phloem in roots, stems, and leaves.
Define meristematic tissue with at least two distinctive features.	BIO-BOT-DOKU-001.1	Meristematic tissue is the embryonic growth tissue composed of undifferentiated, mitotically active, thin-walled, dense-cytoplasm cells that provides longitudinal/lateral growth in the plant.
Write two basic differences between collenchyma and sclerenchyma.	BIO-BOT-DOKU-001.2	Collenchyma cells are living and have an irregularly thickened, non-lignified primary wall (flexible support); sclerenchyma cells are mostly dead and carry a thick, lignified secondary wall (rigid support).

Note. Adapted from the expected core answers of the short-answer items in the tissue module.

So that educators can directly use it for their own concepts, a plain and reusable template for EduCode records is presented in Table 16. The template can also be filled in printed form under offline conditions.

Table 16 Plain EduCode Concept Record Template for Educators (Blank)

Field	Description / fill-in guide (example)
Concept code	Hierarchical identity: field–subfield–concept–version (e.g., BIO-BOT-FOTO-001)
Concept name	Name in the local language (e.g., Fotosentez)
Universal name	International/English equivalent (e.g., Photosynthesis)
Generalization definition (core)	Inclusive, indisputable common-denominator definition
Uniqueness (fingerprint)	Distinctive feature(s) that separate the concept from similar ones
Related concepts	Prerequisite and connected concepts
Misunderstanding risk	Common misconception and EduCode correction
Measurement item	Clarity-indexed short-answer question and the expected core answer

Note. The template is a simplified, copy-and-fill form of the EduCode Protocol (§4.2) for educators.