



Σχολή Θετικών Επιστημών και Τεχνολογίας
Μεταπτυχιακή Εξειδίκευση Καθηγητών των Φυσικών
Επιστημών

Διπλωματική Εργασία

Ph-D A.I. (PhysicsDialog AI): Σχεδιασμός και Αξιολόγηση ενός
Διαλογικού Συστήματος Τεχνητής Νοημοσύνης για την Ανίχνευση
Παρανοήσεων στη Διδασκαλία της Φυσικής

Γιώργος Βαλαβανίδης

Επιβλέπων Καθηγητής: Σταύρος Νόνης

Πάτρα, Οκτώβριος 2026

Η παρούσα εργασία αποτελεί πνευματική ιδιοκτησία του φοιτητή («συγγραφέας/δημιουργός») που την εκπόνησε. Στο πλαίσιο της πολιτικής ανοικτής πρόσβασης ο συγγραφέας/δημιουργός εκχωρεί στο ΕΑΠ, μη αποκλειστική άδεια χρήσης του δικαιώματος αναπαραγωγής, προσαρμογής, δημόσιου δανεισμού, παρουσίας στο κοινό και ψηφιακής διάχυσής τους διεθνώς, σε ηλεκτρονική μορφή και σε οποιοδήποτε μέσο, για διδακτικούς και ερευνητικούς σκοπούς, άνευ ανταλλάγματος και για όλο το χρόνο διάρκειας των δικαιωμάτων πνευματικής ιδιοκτησίας. Η ανοικτή πρόσβαση στο πλήρες κείμενο για μελέτη και ανάγνωση δεν σημαίνει καθ' οιονδήποτε τρόπο παραχώρηση δικαιωμάτων διανοητικής ιδιοκτησίας του συγγραφέα/δημιουργού ούτε επιτρέπει την αναπαραγωγή, αναδημοσίευση, αντιγραφή, αποθήκευση, πώληση, εμπορική χρήση, μετάδοση, διανομή, έκδοση, εκτέλεση, «μεταφόρτωση» (downloading), «ανάρτηση» (uploading), μετάφραση, τροποποίηση με οποιονδήποτε τρόπο, τμηματικά ή περιληπτικά της εργασίας, χωρίς τη ρητή προηγούμενη έγγραφη συναίνεση του συγγραφέα/δημιουργού. Ο συγγραφέας/δημιουργός διατηρεί το σύνολο των ηθικών και περιουσιακών του δικαιωμάτων.



Ph-D A.I. (PhysicsDialog AI): Σχεδιασμός και Αξιολόγηση ενός
Διαλογικού Συστήματος Τεχνητής Νοημοσύνης για την Ανίχνευση
Παρανοήσεων στη Διδασκαλία της Φυσικής

Γιώργος Βαλαβανίδης

Επιτροπή Επίβλεψης Διπλωματικής Εργασίας

Επιβλέπων Καθηγητής:

Σταύρος Νόνης

ΣΕΠ, Ελληνικό Ανοικτό Πανεπιστήμιο

Πάτρα, Οκτώβριος 2026

Περίληψη

Η παρούσα διπλωματική εργασία εξετάζει τη λειτουργία του διαλογικού συστήματος τεχνητής νοημοσύνης Ph-D A.I. (PhysicsDialog AI) στην ενότητα «Μάζα και Βάρος» του Γυμνασίου. Το σύστημα αναπτύχθηκε στο πλαίσιο της εργασίας και παρέχεται μέσω της πλατφόρμας imathia-edu.gr. Συνδυάζει αρχές διαλογικής διδασκαλίας, τη φόρτωση επιλεγμένων σχολικών σημειώσεων στο γλωσσικό μοντέλο ως απλοποιημένη, τοπική προσέγγιση Ανάκτησης-Ενισχυμένης Παραγωγής (Retrieval-Augmented Generation, RAG), χωρίς σημασιολογική ή διανυσματική αναζήτηση, και το προτεινόμενο πλαίσιο αξιολόγησης VCLRCI-Q. Το πλαίσιο περιλαμβάνει είκοσι ελέγχους ανά αιτιολόγηση, πέντε για καθέναν από τους δείκτες Validity (Εγκυρότητα), Logic (Λογική), Regulation (Εκπαιδευτική καταλληλότητα) και Clarity (Σαφήνεια), καθώς και πέντε ελέγχους Consistency (Συνέπεια), οι οποίοι εφαρμόζονται στο τέλος του κουίζ.

Η μελέτη ακολουθεί διερευνητική προσέγγιση μελέτης περίπτωσης και εξετάζει τρία συνθετικά σενάρια που εκτελέστηκαν στην εφαρμογή. Τα δύο ενσωματώνουν προκαθορισμένες παρανοήσεις σχετικές με τη μάζα και το βάρος, ενώ το τρίτο λειτουργεί ως σενάριο αναφοράς. Οι αιτιολογήσεις διαμορφώθηκαν και εισήχθησαν από τον ερευνητή ως συνθετικές αλληλεπιδράσεις. Η ανάλυση βασίστηκε σε χειροκίνητα αντίγραφα των διαλόγων, στις βαθμολογίες κάθε ερώτησης και στις τελικές αξιολογήσεις του συστήματος.

Ο μέσος συνολικός δείκτης Q υπολογίστηκε σε 0,35, 0,36 και 1,00, σε συμφωνία με τις τιμές του συστήματος. Η ανάλυση ανέδειξε ότι η σαφής διατύπωση μπορεί να συνυπάρχει με επιστημονικά εσφαλμένη αιτιολόγηση. Ο δείκτης συνέπειας C ήταν 0,40 στα δύο πρώτα σενάρια και 1,00 στο σενάριο αναφοράς. Στο Σενάριο 2 η τιμή αντιστοιχεί σε συγκεκριμένη αντίφαση μεταξύ ισχυρισμών, ενώ στο Σενάριο 1 αποτυπώνει τη συνολική κρίση του μοντέλου για το κουίζ. Η ανατροφοδότηση ακολούθησε τη σχολική διάκριση μάζας και βάρους και η διαλογική ακολουθία των τριών ρόλων τηρήθηκε σε όλες τις ερωτήσεις.

Η συμβολή της εργασίας εστιάζει στον σχεδιασμό και στη λειτουργική διερεύνηση ενός συστήματος που συνδυάζει αιτιολόγηση πριν από την ανατροφοδότηση, τρεις παιδαγωγικούς ρόλους και πολυδιάστατη περιγραφική αξιολόγηση. Τα ευρήματα τεκμηριώνουν τη λειτουργία του συστήματος στα συγκεκριμένα συνθετικά σενάρια και διαμορφώνουν τη βάση για επόμενη εμπειρική αξιολόγηση. Ως επόμενα βήματα προτείνονται η εφαρμογή με

πραγματικούς μαθητές, η ανεξάρτητη αξιολόγηση των ίδιων απαντήσεων από ειδικούς και η διερεύνηση της αξιοπιστίας και της εγκυρότητας του πλαισίου.

Λέξεις – Κλειδιά

διαλογική διδασκαλία, παρανοήσεις στη Φυσική, μάζα και βάρος, μεγάλα γλωσσικά μοντέλα, ευφυή συστήματα διδασκαλίας, ανατροφοδότηση

Ph-D A.I. (PhysicsDialog AI): Design and Evaluation of a Dialogic Artificial Intelligence System for Detecting Misconceptions in Physics Teaching

Giorgos Valavanidis

Abstract

This thesis examines the operation of the dialogic artificial intelligence system Ph-D A.I. (PhysicsDialog AI) in the lower-secondary instructional unit "Mass and Weight." The system was developed as part of the study and is available through the imathia-edu.gr platform. It combines principles of dialogic teaching, the loading of selected school notes into the language model as a simplified, local Retrieval-Augmented Generation (RAG) approach without semantic or vector-based retrieval, and the proposed VCLRCI-Q assessment framework. The framework includes twenty checks per explanation, five for each of the indicators Validity, Logic, Regulation (suitability for the educational level), and Clarity, as well as five Consistency checks applied at the end of the quiz.

The study follows an exploratory case-study approach and examines three synthetic scenarios executed in the application. Two incorporate predefined misconceptions concerning mass and weight, while the third serves as a reference scenario. The explanations were designed and entered by the researcher as synthetic interactions. The analysis was based on manually copied dialogues, the scores of each question, and the system's final assessments.

The mean Q scores were calculated as 0.35, 0.36, and 1.00, respectively, in agreement with the values reported by the system. The analysis showed that clear expression can coexist with scientifically incorrect reasoning. The Consistency score C was 0.40 in the first two scenarios and 1.00 in the reference scenario. In Scenario 2, the value corresponds to a specific contradiction between claims, whereas in Scenario 1 it reflects the model's overall judgment of the quiz. Feedback followed the school-level distinction between mass and weight, and the three-role dialogue sequence was maintained in every question.

The contribution of the study lies in the design and functional exploration of a system that combines explanation before feedback, three pedagogical roles, and multidimensional descriptive assessment. The findings document the operation of the system in the specific synthetic scenarios and provide a basis for subsequent empirical evaluation. Next steps

include implementation with real students, independent expert assessment of the same responses, and investigation of the reliability and validity of the framework.

Keywords

dialogic teaching, physics misconceptions, mass and weight, large language models, intelligent tutoring systems, feedback

Περιεχόμενα

Περίληψη.....	iv
Abstract	vi
Περιεχόμενα	viii
Κατάλογος Πινάκων	xii
Συντομογραφίες & Ακρωνύμια.....	xiv
1. Εισαγωγή.....	1
1.1 Αντικείμενο και Σκοπός της Εργασίας	1
1.2 Ερευνητικά Ερωτήματα	2
1.3 Δομή της Εργασίας.....	2
1.4 Συμβολή της Εργασίας.....	3
1.5 Πεδίο της Μελέτης	4
2. Θεωρητικό Πλαίσιο	5
2.1 Διαλογική Διδασκαλία στις Φυσικές Επιστήμες	5
2.1.1 Ερωτήσεις ως Εργαλείο Διαλογικής Μάθησης	6
2.1.2 Σωκρατικός Διάλογος και Τριπλός Ρόλος	7
2.2 Παρανοήσεις στη Διδασκαλία της Φυσικής	7
2.2.1 Ταξινόμηση Παρανοήσεων	8
2.2.2 Παρανοήσεις στην Ενότητα Μάζα-Βάρος	8
2.2.3 Θεωρία Εννοιολογικής Αλλαγής.....	9
2.3 Τεχνητή Νοημοσύνη ως Παιδαγωγικό Εργαλείο	9
2.3.1 Από τα Παραδοσιακά ITS στα Σύγχρονα LLM	10
2.4 Ποιότητα Ανατροφοδότησης και Μάθηση	12
2.4.1 Το Μοντέλο Hattie & Timperley (2007)	12
2.4.2 Επεξηγηματική Ανατροφοδότηση (Elaborated Feedback).....	13
2.4.3 Χρονισμός και Προσαρμογή	13
2.5 Το Πλαίσιο Αξιολόγησης VCLRCI-Q — Θεωρητική Θεμελίωση	14
2.6 Σύνοψη Θεωρητικού Πλαισίου	15
3. Το Σύστημα Ph-D A.I. (PhysicsDialog AI)	16
3.1 Αρχιτεκτονική Συστήματος	16
3.1.1 Frontend — Διεπαφή Χρήστη	17
3.1.2 Διακομιστής Εφαρμογής (Backend – PHP)	17
3.1.3 Γλωσσικό Μοντέλο (Claude Haiku 4.5).....	18
3.2 RAG (Retrieval-Augmented Generation) — Ανάκτηση-Ενισχυμένη Παραγωγή.....	19
3.2.1 Δομή Αρχείων RAG	19

3.2.2 Ενσωμάτωση του Γνωστικού Υλικού στο Prompt.....	19
3.2.3 Σχεδιαστικές δυνατότητες της παροχής σημειώσεων	20
3.3 Prompt Caching (Προσωρινή Αποθήκευση Οδηγιών)	20
3.4 Προσαρμοστική Δυσκολία (Adaptive Difficulty)	21
3.5 Διαλογική Δομή — Οι Τρεις Ρόλοι	22
3.6 Κύκλος Λειτουργίας	23
3.7 Ρουμπρίκα Αξιολόγησης VCLRCI-Q	24
3.7.1 Οι 20 έλεγχοι ανά αιτιολόγηση	24
3.7.2 Τα C Checks — Consistency στο Τέλος	25
3.7.3 Υπολογισμός Δεικτών Αξιολόγησης	26
3.7.4 Ο Συνολικός Δείκτης Q	26
3.7.5 Επιλογή Συντελεστών Βαρύτητας	27
3.8 Δομημένα Σήματα Καταγραφής (Backend Signals)	27
3.8.1 Δομημένη Καταγραφή Αξιολόγησης (Q_RECORD) Ανά Ερώτηση	27
3.8.2 Σύνοψη Συνεδρίας (QUIZ_SUMMARY)	28
3.9 Μηχανισμός Επανεξήγησης (Re-explain Flow)	28
3.10 Αποθήκευση Δεδομένων	29
3.11 Τεχνικοί Περιορισμοί του Συστήματος	30
3.12 Σύνοψη Κεφαλαίου	30
4. Μεθοδολογία Έρευνας.....	31
4.1 Ερευνητικά Ερωτήματα	31
4.2 Ερευνητικός Σχεδιασμός — Μελέτη Περίπτωσης.....	32
4.3 Σχεδιασμός Συνθετικών Σεναρίων	33
4.3.1 Επιλογή Παρανοήσεων από Βιβλιογραφία	33
4.3.2 Σχεδιασμός Απαντήσεων	33
4.3.3 Εκτέλεση στο Πραγματικό Σύστημα	34
4.4 Εργαλεία Συλλογής Δεδομένων	34
4.4.1 Χειροκίνητη Καταγραφή Ορατού Διαλόγου	34
4.4.2 Δεδομένα ανά Ερώτηση (Q_RECORD).....	35
4.4.3 Συνοπτικά Δεδομένα Συνεδρίας (QUIZ_SUMMARY).....	35
4.5 Διαδικασία Ανάλυσης Δεδομένων	35
4.5.1 Πρώτο Επίπεδο — Ποιοτική Ανάλυση Διαλόγων.....	35
4.5.2 Δεύτερο Επίπεδο — Ανάλυση Δεικτών	36
4.5.3 Τρίτο Επίπεδο — Συγκριτική Ανάλυση.....	36
4.6 Μεθοδολογικοί Περιορισμοί	37
4.7 Ηθικά και Δεοντολογικά Ζητήματα	38

4.8 Αξιοπιστία και Εγκυρότητα	39
4.9 Καθεστώς Ερευνητικής Χρήσης των Δεικτών	39
4.10 Σύνοψη Μεθοδολογίας	40
5. Ανάλυση Δεδομένων και Αποτελεσμάτων.....	41
5.1 Πλαίσιο Σεναρίων	41
5.2 Σενάριο 1 (Μαθητής_A)	41
5.2.1 Ανάλυση ανά Ερώτηση.....	41
5.2.2 Τιμές Δεικτών ανά Ερώτηση Σεναρίου 1	44
5.2.3 C Score — Consistency Analysis (Ανάλυση Συνέπειας).....	45
5.2.4 Τελικό Q Score Σεναρίου 1	45
5.2.5 Αξιολόγηση Ανατροφοδότησης.....	46
5.3 Σενάριο 2 (Μαθητής_B)	46
5.3.1 Ανάλυση ανά Ερώτηση.....	47
5.3.2 Τιμές Δεικτών ανά Ερώτηση Σεναρίου 2	49
5.3.3 C Score — Consistency Analysis.....	50
5.3.4 Τελικό Q Score Σεναρίου 2	51
5.4 Σενάριο 3 (Μαθητής_Γ).....	51
5.4.1 Score Analysis Αναφοράς	51
5.4.2 C Score Αναφοράς.....	52
5.5 Συγκριτική Ανάλυση Σεναρίων	53
5.5.1 Σύγκριση Δεικτών	53
5.5.2 Ευρήματα ανά Ερευνητικό Ερώτημα	53
5.5.3 Κύρια Ευρήματα	54
5.6 Προέλευση Δεδομένων και Έλεγχος Αριθμητικής Συμφωνίας	55
5.7 Περιορισμοί Ανάλυσης.....	56
5.8 Σύνοψη Κεφαλαίου	56
6. Συμπεράσματα και Προτάσεις.....	57
6.1 Κύρια Συμπεράσματα	57
6.1.1 Ευρήματα από την Ανάλυση Σεναρίων	57
6.1.2 Σύνοψη Δεικτών	57
6.2 Απαντήσεις στα Ερευνητικά Ερωτήματα	58
6.3 Θεωρητική Συμβολή της Εργασίας.....	59
6.3.1 Στη Διδακτική Φυσικών Επιστημών	59
6.3.2 Στην Αξιολόγηση και Ανατροφοδότηση	59
6.3.3 Στην Τεχνητή Νοημοσύνη στην Εκπαίδευση.....	60
6.4 Παιδαγωγικές Προεκτάσεις	60

6.4.1 Για τον Εκπαιδευτικό	60
6.4.2 Αρχές Σχεδιασμού Εκπαιδευτικών Συστημάτων Τεχνητής Νοημοσύνης	60
6.5 Περιορισμοί της Έρευνας	61
6.6 Προοπτικές Περαιτέρω Διερεύνησης	61
6.6.1 Εμπειρική Αξιολόγηση	61
6.6.2 Επικύρωση Δεικτών	62
6.6.3 Επέκταση Περιεχομένου	63
6.6.4 Τεχνολογικές Βελτιώσεις	63
6.7 Συνοπτική Αποτίμηση της Εργασίας	64
Βιβλιογραφία	65
Παραρτήματα — Διάλογοι των Συνθετικών Σεναρίων	67
Παράρτημα Α : Σενάριο 1 (Μαθητής_A)	67
Παράρτημα Β — Σενάριο 2 (Μαθητής_B)	73
Παράρτημα Γ : Σενάριο 3 (Μαθητής_Γ)	79
Παράρτημα Δ : Δομές Q_RECORD και QUIZ_SUMMARY	85

Κατάλογος Πινάκων

Πίνακας 2.1: Χαρακτηριστικά Διαλογικής Διδασκαλίας (Alexander, 2008) και Εφαρμογή τους	6
Πίνακας 2.2: Τύποι Ερωτήσεων και Εφαρμογή τους στο Ph-D A.I. (PhysicsDialog AI).....	6
Πίνακας 2.3: Κατηγορίες Παρανοήσεων	8
Πίνακας 2.4: Ενδεικτικές Παρανοήσεις στην Ενότητα Μάζα-Βάρος	9
Πίνακας 2.5: Σύγκριση ITS με Ρητά Μοντέλα και Συστημάτων με LLM	11
Πίνακας 2.6: Ενδεικτικές Προσεγγίσεις Εκπαιδευτικών Συστημάτων	12
Πίνακας 3.1: Αρχιτεκτονικά Στρώματα του Ph-D A.I. (PhysicsDialog AI)	17
Πίνακας 3.2: Παροχή Σημειώσεων και Γενική Γνώση LLM	20
Πίνακας 3.3: Οι Τρεις Ρόλοι του Συστήματος	22
Πίνακας 3.4: Οι 20 Έλεγχοι ανά Αιτιολόγηση.....	25
Πίνακας 3.5: Τα 5 C Checks για Consistency	25
Πίνακας 3.6: Περιγραφικοί Χαρακτηρισμοί του Τελικού Q	26
Πίνακας 3.7: Τεχνικοί Περιορισμοί και Μετριάστικοι Παράγοντες	30
Πίνακας 4.1: Ερευνητικά Ερωτήματα, Πηγές και Τρόπος Ανάλυσης	32
Πίνακας 4.2: Παράμετροι Συγκριτικής Ανάλυσης.....	37
Πίνακας 4.3: Μεθοδολογικοί Περιορισμοί και Επιπτώσεις.....	38
Πίνακας 5.1: Score Analysis (Ανάλυση Βαθμολογίας) Ερ.1 Σεναρίου 1	42
Πίνακας 5.2: Score Analysis Ερ.2 Σεναρίου 1	42
Πίνακας 5.3: Score Analysis Ερ.3 Σεναρίου 1	43
Πίνακας 5.4: Score Analysis Ερ.4 Σεναρίου 1	43
Πίνακας 5.5: Score Analysis Ερ.5 Σεναρίου 1	44
Πίνακας 5.6: Τιμές Δεικτών Σεναρίου 1	44
Πίνακας 5.7: C Checks Σεναρίου 1	45
Πίνακας 5.8: Υπολογισμός Q Score Σεναρίου 1	46
Πίνακας 5.9: Τύπος Ανατροφοδότησης ανά Ερώτηση — Σενάριο 1	46
Πίνακας 5.10: Score Analysis Ερ.1 Σεναρίου 2.....	47
Πίνακας 5.11: Score Analysis Ερ.2 Σεναρίου 2	48
Πίνακας 5.12: Score Analysis Ερ.3 Σεναρίου 2.....	48
Πίνακας 5.13: Score Analysis Ερ.4 Σεναρίου 2.....	49
Πίνακας 5.14: Score Analysis Ερ.5 Σεναρίου 2.....	49
Πίνακας 5.15: Τιμές Δεικτών Σεναρίου 2	50
Πίνακας 5.16: C Checks Σεναρίου 2	50
Πίνακας 5.17: Υπολογισμός Q Score Σεναρίου 2	51

Πίνακας 5.18: Score Analysis Σεναρίου 3	52
Πίνακας 5.19: C Checks Σεναρίου 3	52
Πίνακας 5.20: Συγκριτική Σύνοψη Δεικτών	53
Πίνακας 5.21: Κύρια Ευρήματα	55
Πίνακας 6.1: Συνοπτική Συμπεριφορά Δεικτών στα Τρία Σενάρια	58

Συντομογραφίες & Ακρωνύμια

AI	Artificial Intelligence (Τεχνητή Νοημοσύνη)
AIED	Artificial Intelligence in Education (Τεχνητή Νοημοσύνη στην Εκπαίδευση)
API	Application Programming Interface (Διεπαφή Προγραμματισμού Εφαρμογών)
C	Consistency (Συνέπεια)
CSS	Cascading Style Sheets
HTML	HyperText Markup Language
ITS	Intelligent Tutoring Systems (Ευφυή Συστήματα Διδασκαλίας)
JSON	JavaScript Object Notation
L	Logic (Λογική)
LLM	Large Language Model (Μεγάλο Γλωσσικό Μοντέλο)
Ph-D A.I.	PhysicsDialog AI
PHP	PHP: Hypertext Preprocessor
Q	Συνολικός δείκτης αξιολόγησης της ρουμπρίκας VCLRCI-Q
R	Regulation (Εκπαιδευτική καταλληλότητα)
RAG	Retrieval-Augmented Generation (Ανάκτηση-Ενισχυμένη Παραγωγή)
V	Validity (Εγκυρότητα)
VCLRCI-Q	Validity, Consistency, Logic, Regulation, Clarity – Q (πλαίσιο αξιολόγησης)
ΕΑΠ	Ελληνικό Ανοικτό Πανεπιστήμιο
ΕΕ	Ερευνητικό Ερώτημα
ΣΕΠ	Συνεργαζόμενο Εκπαιδευτικό Προσωπικό
ΣΘΕΤ	Σχολή Θετικών Επιστημών και Τεχνολογίας

1. Εισαγωγή

1.1 Αντικείμενο και Σκοπός της Εργασίας

Η ενσωμάτωση ψηφιακών τεχνολογιών στη σύγχρονη εκπαιδευτική πραγματικότητα έχει μετασχηματίσει ριζικά τον τρόπο με τον οποίο σχεδιάζεται, υλοποιείται και αξιολογείται η μαθησιακή διαδικασία (Luckin et al., 2016). Παρά την εκτεταμένη χρήση ψηφιακών εργαλείων και πλατφορμών, η διδασκαλία των Φυσικών Επιστημών εξακολουθεί να παρουσιάζει ιδιαίτερες δυσκολίες, κυρίως λόγω του αφηρημένου χαρακτήρα πολλών εννοιών και της συχνής εμφάνισης παρανοήσεων που εδραιώνονται στη σκέψη των μαθητών.

Στο γνωστικό πεδίο της Φυσικής, έννοιες όπως η μάζα, η δύναμη και η τριβή είναι χαρακτηριστικά παραδείγματα εννοιολογικών εμποδίων. Οι μαθητές συχνά διαμορφώνουν εναλλακτικές ερμηνείες που παραμένουν ανθεκτικές, ακόμη και μετά από παραδοσιακή διδασκαλία (Driver et al., 1985· Viennot, 1979). Το γεγονός αυτό καθιστά αναγκαία την ανάπτυξη παιδαγωγικών προσεγγίσεων που δεν περιορίζονται στην ορθότητα των απαντήσεων, αλλά εστιάζουν στη διερεύνηση της σκέψης και των συλλογιστικών μοτίβων του μαθητή.

Στο πλαίσιο αυτό, η παρούσα διπλωματική εργασία παρουσιάζει τον σχεδιασμό και τη λειτουργική διερεύνηση ενός διαλογικού συστήματος τεχνητής νοημοσύνης για τη διδασκαλία της Φυσικής. Το σύστημα, με την ονομασία Ph-D A.I. (PhysicsDialog AI), αναπτύχθηκε από τον συγγραφέα ως διαδικτυακή εφαρμογή στη διεύθυνση imathia-edu.gr. Συνδυάζει αρχές διαλογικής διδασκαλίας (Alexander, 2008), τη φόρτωση επιλεγμένων σημειώσεων του κεφαλαίου στο γλωσσικό μοντέλο, δομημένη αξιολόγηση της μαθητικής αιτιολόγησης και οδηγίες για προσαρμοσμένη ανατροφοδότηση (Shute, 2008). Η φόρτωση των σημειώσεων αποτελεί μια απλοποιημένη μορφή Ανάκτησης-Ενισχυμένης Παραγωγής (Retrieval-Augmented Generation, RAG): το σύστημα ανακτά τις σημειώσεις του κεφαλαίου που επιλέγει ο μαθητής, χωρίς σημασιολογική αναζήτηση.

Ο βασικός σκοπός είναι να διερευνηθεί η λειτουργία του συστήματος σε τρία συνθετικά σενάρια ως προς:

1. την αναγνώριση προκαθορισμένων παρανοήσεων μέσω της ανάλυσης της επιλογής και της αιτιολόγησης των απαντήσεων,

2. την παραγωγή ανατροφοδότησης προσαρμοσμένης στον συλλογισμό που διατυπώνεται, και
3. την εφαρμογή δομημένων κριτηρίων αξιολόγησης και τη διαλογική επεξεργασία των φυσικών εννοιών μέσω των τριών παιδαγωγικών ρόλων.

Το Ph-D A.I. (PhysicsDialog AI) σχεδιάστηκε ως υποστηρικτικό παιδαγωγικό εργαλείο για τον εκπαιδευτικό, με στόχο τη διερεύνηση αιτιολογήσεων και εννοιολογικών δυσκολιών. Η παρούσα μελέτη εστιάζει στη λειτουργική διερεύνηση του συστήματος μέσω συνθετικών αλληλεπιδράσεων, δηλαδή διαλόγων με απαντήσεις που διαμόρφωσε ο ερευνητής.

1.2 Ερευνητικά Ερωτήματα

Με βάση το παραπάνω πλαίσιο, η εργασία εξετάζει τα ακόλουθα ερευνητικά ερωτήματα:

EE1: Με ποιον τρόπο το διαλογικό σύστημα αναγνωρίζει και αποτυπώνει προκαθορισμένες παρανοήσεις στις αιτιολογήσεις των συνθετικών σεναρίων;

EE2: Ποια χαρακτηριστικά εμφανίζει η παραγόμενη ανατροφοδότηση και πώς συνδέονται με τις αρχές της εκπαιδευτικής ανατροφοδότησης των Hattie και Timperley (2007);

EE3: Πώς αξιοποιούνται οι δείκτες Validity (V, εγκυρότητα), Consistency (C, συνέπεια), Logic (L, λογική), Regulation (R, εκπαιδευτική καταλληλότητα) και Clarity (Cl, σαφήνεια) για την περιγραφική και συγκριτική ανάλυση των απαντήσεων και της αιτιολόγησης στα τρία συνθετικά σενάρια;

EE4: Πώς οργανώνεται η διαλογική αλληλεπίδραση μέσω των τριών παιδαγωγικών ρόλων και ποιες δυνατότητες αιτιολόγησης και αναστοχασμού δημιουργεί ο σχεδιασμός της;

1.3 Δομή της Εργασίας

Η εργασία αναπτύσσεται σε έξι κεφάλαια, τα οποία ακολουθούν την πορεία από τη θεωρητική θεμελίωση έως την εφαρμογή και τα τελικά συμπεράσματα.

Κεφάλαιο 1 (παρόν): Παρουσιάζει το αντικείμενο, τον σκοπό, τα ερευνητικά ερωτήματα και τη δομή της εργασίας.

Κεφάλαιο 2: Αναπτύσσει το θεωρητικό πλαίσιο σε τρεις άξονες: (α) διαλογική διδασκαλία, (β) παρανοήσεις στη Φυσική, (γ) τεχνητή νοημοσύνη ως παιδαγωγικό εργαλείο.

Κεφάλαιο 3: Περιγράφει την τεχνική αρχιτεκτονική του Ph-D A.I. (PhysicsDialog AI), τη φόρτωση σημειώσεων κεφαλαίου, το prompt caching (προσωρινή αποθήκευση των οδηγιών του μοντέλου) και τους τρεις ρόλους διαλόγου. Παρουσιάζει επίσης τη ρουμπρίκα VCLRCI-Q, η οποία περιλαμβάνει 20 διακριτούς ελέγχους ανά μαθητική αιτιολόγηση και πέντε επιπλέον ελέγχους συνέπειας στο τέλος της συνεδρίας. Τέλος, περιγράφονται ο υπολογισμός του δείκτη Q, τα δομημένα σήματα καταγραφής και οι μηχανισμοί διαχείρισης δεδομένων.

Κεφάλαιο 4: Παρουσιάζει τη μεθοδολογία, τον σχεδιασμό και την εκτέλεση τριών συνθετικών σεναρίων, τις πηγές δεδομένων (χειροκίνητα καταγραμμένοι διάλογοι, ορατές τιμές δεικτών και διαθέσιμες τελικές συνόψεις), τη διαδικασία ανάλυσης, τη διάκριση των δεδομένων που αναλύθηκαν από τις εγγραφές Q_RECORD (δομημένες εγγραφές αξιολόγησης ανά ερώτηση, που δεν εμφανίζονται στον χρήστη) και τους μεθοδολογικούς περιορισμούς.

Κεφάλαιο 5: Παρουσιάζει την ανάλυση των τριών σεναρίων — τις τιμές των δεικτών ανά ερώτηση, τον έλεγχο του υπολογισμού του Q, τη συγκριτική ανάλυση, τα κύρια ευρήματα και τις απαντήσεις στα ερευνητικά ερωτήματα.

Κεφάλαιο 6: Συνοψίζει τα συμπεράσματα, τις παιδαγωγικές προεκτάσεις, τους περιορισμούς και τις προτάσεις μελλοντικής έρευνας.

1.4 Συμβολή της Εργασίας

Η συμβολή της παρούσας εργασίας αναπτύσσεται σε τρεις αλληλένδετους άξονες: τον παιδαγωγικό, τον τεχνολογικό και τον ερευνητικό.

Παιδαγωγικός άξονας

Προτείνεται ένα μοντέλο ποιοτικής ανατροφοδότησης που εστιάζει στη δομή και στη συνέπεια της μαθητικής αιτιολόγησης και όχι αποκλειστικά στην ορθότητα των απαντήσεων. Η προσέγγιση αυτή στηρίζεται στις αρχές της εκπαιδευτικής ανατροφοδότησης (Hattie & Timperley, 2007) και της διαλογικής διδασκαλίας (Alexander, 2008).

Τεχνολογικός άξονας

Αναπτύχθηκε ένα διαλογικό εκπαιδευτικό σύστημα που συνδυάζει επιλεγμένο εκπαιδευτικό υλικό, πολυδιάστατη αξιολόγηση της μαθητικής αιτιολόγησης και προσαρμοσμένη ανατροφοδότηση. Το σύστημα σχεδιάστηκε με έμφαση στη μαθησιακή διαδικασία και όχι μόνο στην παραγωγή απαντήσεων.

Ερευνητικός άξονας

Στο πλαίσιο της μελέτης διαμορφώθηκε η ρουμπρίκα VCLRCI-Q και εξετάστηκε η λειτουργία της στα τρία συνθετικά σενάρια. Οι δείκτες της περιγράφουν την επιστημονική ορθότητα, τη συνέπεια μεταξύ απαντήσεων, τη λογική ακολουθία, την καταλληλότητα της αιτιολόγησης για το εκπαιδευτικό επίπεδο και τη σαφήνεια της διατύπωσης. Τα κριτήρια περιλαμβάνονται στο system prompt (τις βασικές οδηγίες που λαμβάνει το γλωσσικό μοντέλο) και τα εφαρμόζει το ίδιο το μοντέλο. Οι τιμές των δεικτών αξιοποιούνται ως λειτουργικές ενδείξεις για την περιγραφική και συγκριτική ανάλυση των συγκεκριμένων σεναρίων. Η εμπειρική επικύρωσή τους ως δεικτών μαθητικής επίδοσης ή γνωστικής προόδου, καθώς και η σύγκριση των κρίσεών τους με ανεξάρτητους αξιολογητές, αποτελούν αντικείμενο μελλοντικής έρευνας.

1.5 Πεδίο της Μελέτης

Η μελέτη αφορά την ενότητα «Μάζα και Βάρος» της σχολικής Φυσικής και συνθετικά προφίλ απαντήσεων για τη Β΄ Γυμνασίου. Η επιλογή του γνωστικού πεδίου στηρίζεται στη βιβλιογραφία για τις παρανοήσεις σχετικά με τη μάζα και το βάρος (Driver et al., 1985· Viennot, 1979· Galili, 1993).

Η έρευνα βασίστηκε σε τρία συνθετικά σενάρια αλληλεπίδρασης που εκτελέστηκαν στο λειτουργικό περιβάλλον της εφαρμογής. Τα δύο περιλάμβαναν προκαθορισμένες παρανοήσεις, ενώ το τρίτο λειτουργούσε ως σενάριο αναφοράς με επιστημονικά ορθές και συνεπείς απαντήσεις. Κάθε σενάριο περιλάμβανε πέντε ερωτήσεις. Δεν συμμετείχαν πραγματικοί μαθητές.

Η ανάλυση εστιάζει στις αποκρίσεις του συστήματος, στην εφαρμογή της ρουμπρίκας και στη διαφοροποίηση των αποτελεσμάτων μεταξύ των τριών συνθετικών σεναρίων. Ως εκ τούτου, τα συμπεράσματα αφορούν τη λειτουργία του συστήματος στο συγκεκριμένο ερευνητικό πλαίσιο, ενώ η αξιολόγηση με πραγματικούς μαθητές και η διερεύνηση μαθησιακών αποτελεσμάτων αποτελούν επόμενο στάδιο εμπειρικής έρευνας. Η μεθοδολογική οριοθέτηση παρουσιάζεται αναλυτικά στο Κεφάλαιο 4 και οι συνολικοί περιορισμοί εξετάζονται στο Κεφάλαιο 6.

2. Θεωρητικό Πλαίσιο

Η παρούσα διπλωματική εργασία εντάσσεται στη διδακτική των Φυσικών Επιστημών και εξετάζει τον σχεδιασμό και τη λειτουργία ενός διαλογικού ψηφιακού περιβάλλοντος για τη Φυσική. Το θεωρητικό υπόβαθρο εξηγεί πώς η αιτιολόγηση, η εξέταση των αρχικών αντιλήψεων και η επεξηγηματική ανατροφοδότηση μπορούν να υποστηρίξουν την εννοιολογική κατανόηση, και αποτελεί τη βάση των παιδαγωγικών επιλογών του σχεδιασμού.

Το θεωρητικό πλαίσιο οργανώνεται γύρω από τρεις αλληλένδετους άξονες: (α) τη διαλογική διδασκαλία, (β) τη θεωρία των παρανοήσεων στη Φυσική και (γ) τον ρόλο της τεχνητής νοημοσύνης ως εργαλείου παιδαγωγικής ανατροφοδότησης. Οι άξονες αυτοί συγκροτούν το παιδαγωγικό υπόβαθρο πάνω στο οποίο σχεδιάστηκε το Ph-D A.I. (PhysicsDialog AI).

2.1 Διαλογική Διδασκαλία στις Φυσικές Επιστήμες

Η διαλογική προσέγγιση στη διδασκαλία επιδιώκει την ενεργό συμμετοχή και τη γνωστική αλληλεπίδραση. Η συζήτηση, η ανταλλαγή επιχειρημάτων και η διαπραγμάτευση νοήματος προσφέρουν ευκαιρίες εξέτασης των ιδεών των μαθητών και του διδάσκοντος. Στο Ph-D A.I. οι αρχές αυτές αξιοποιούνται ως οδηγός σχεδιασμού μιας προσομοιωμένης διαλογικής αλληλεπίδρασης.

Σύμφωνα με τον Alexander (2008), η διαλογική διδασκαλία είναι μια μορφή εκπαιδευτικής επικοινωνίας που βασίζεται στη συλλογική διερεύνηση ιδεών. Ο Alexander ορίζει πέντε βασικά χαρακτηριστικά που διακρίνουν τη διαλογική διδασκαλία από την παραδοσιακή:

Χαρακτηριστικό	Περιγραφή	Εφαρμογή στο Ph-D A.I. (PhysicsDialog AI)
Συλλογικό (Collective)	Κοινή διερεύνηση ιδεών και εννοιών	Οι τρεις προσομοιωμένοι ρόλοι παρουσιάζουν διαφορετικές οπτικές για το ίδιο πρόβλημα, μιμούμενοι μια συλλογική διερεύνηση.
Αμοιβαίο (Reciprocal)	Αμοιβαία ανταλλαγή απόψεων και επιχειρημάτων	Ο μαθητής διατυπώνει επιλογή και αιτιολόγηση· ακολουθούν εξήγηση, διάλογος ρόλων και ερώτηση κατανόησης.

Χαρακτηριστικό	Περιγραφή	Εφαρμογή στο Ph-D A.I. (PhysicsDialog AI)
Υποστηρικτικό (Supportive)	Δημιουργία ασφαλούς μαθησιακού περιβάλλοντος	Το prompt (οι οδηγίες προς το γλωσσικό μοντέλο) προβλέπει κατανοητές εξηγήσεις και επανεξήγηση κάθε φορά που ο μαθητής δηλώνει ότι δεν κατάλαβε.
Συσσωρευτικό (Cumulative)	Οι νέες ιδέες οικοδομούνται πάνω στις προηγούμενες	Οι εξηγήσεις συνδέονται με την αιτιολόγηση. Το διαθέσιμο ιστορικό προηγούμενων συνεδριών μπορεί να καθοδηγεί τη δημιουργία νέου κουίζ.
Στοχευμένο (Purposeful)	Ο διάλογος υπηρετεί συγκεκριμένους μαθησιακούς στόχους	Ο διάλογος εστιάζει στις έννοιες του επιλεγμένου κεφαλαίου και στην εξέταση της αιτιολόγησης.

Πίνακας 2.1: Χαρακτηριστικά Διαλογικής Διδασκαλίας (Alexander, 2008) και Εφαρμογή τους

2.1.1 Ερωτήσεις ως Εργαλείο Διαλογικής Μάθησης

Τύπος Ερώτησης	Παράδειγμα	Ρόλος στο Σύστημα
Διευκρινιστική	«Τι εννοείς όταν λες ότι η μάζα μειώνεται;»	Ενδεικτική μορφή διερευνητικού ερωτήματος για την αποσαφήνιση της έννοιας, που μπορεί να προκύψει μέσα στη διαλογική ακολουθία.
Αιτιολογική	«Γιατί πιστεύεις ότι το βάρος είναι σταθερό;»	Η αιτιολόγηση ζητείται μετά από κάθε επιλογή και πριν από την ανακοίνωση της ορθότητάς της.
Ερώτηση διερεύνησης αντίφασης	«Πριν είπες X, τώρα λες Y, πώς εξηγείς αυτό;»	Ενδεικτική παιδαγωγική μορφή. Οι αντιφάσεις μεταξύ απαντήσεων εξετάζονται συνολικά με τον C στο τέλος του κουίζ.
Υποθετική	«Τι θα συνέβαινε αν ήμασταν στον Δία;»	Ο ρόλος της Ελισάβετ εισάγει αμφισβήτηση ή υποθετική περίπτωση, σύμφωνα με τη διαλογική ακολουθία του prompt.

Πίνακας 2.2: Τύποι Ερωτήσεων και Εφαρμογή τους στο Ph-D A.I. (PhysicsDialog AI)

Οι στοχευμένες ερωτήσεις (probing questions) μπορούν να βοηθήσουν στην αποσαφήνιση των ιδεών και στην εξέταση του συλλογισμού. Οι Chin και Osborne (2010) συνδέουν τις ερωτήσεις των μαθητών με την ανάπτυξη επιχειρηματολογίας στις Φυσικές Επιστήμες. Στο Ph-D A.I. απαιτείται αιτιολόγηση μετά την επιλογή απάντησης, ενώ οι παιδαγωγικοί ρόλοι διατυπώνουν διερευνητικές και υποθετικές ερωτήσεις. Ο Πίνακας 2.2 παρουσιάζει ενδεικτικούς τύπους ερωτήσεων. Οι ερωτήσεις αυτές προκύπτουν μέσα από την ενιαία διαλογική ακολουθία του συστήματος, ενώ οι τύποι του πίνακα χρησιμοποιούνται για την ανάλυσή τους.

2.1.2 Σωκρατικός Διάλογος και Τριπλός Ρόλος

Ο διάλογος με τρία πρόσωπα στο Ph-D A.I. (PhysicsDialog AI) αντλεί στοιχεία από τη σωκρατική μέθοδο: ερωτήσεις, αιτιολόγηση και κριτική εξέταση των ιδεών. Ο Κύριος Γιώργος παρέχει εξήγηση, η Ζωή θέτει διερευνητικό ερώτημα και η Ελισάβετ εισάγει αμφισβήτηση ή υποθετική περίπτωση. Οι τρεις ρόλοι παράγονται από το ίδιο γλωσσικό μοντέλο σύμφωνα με τις οδηγίες του system prompt (των βασικών οδηγιών του συστήματος): δεν αποτελούν ανεξάρτητους αξιολογητές. Η ακολουθία είναι προκαθορισμένη και προβλέπει δυνατότητα επανεξήγησης. Κάθε φορά που ο μαθητής δηλώνει ότι δεν κατάλαβε, το σύστημα δίνει μία νέα εξήγηση με απλούστερη διατύπωση και διαφορετικό παράδειγμα, χωρίς νέα βαθμολόγηση. Η τεχνική εφαρμογή παρουσιάζεται στις Ενότητες 3.5 και 3.9.

2.2 Παρανοήσεις στη Διδασκαλία της Φυσικής

Η έρευνα στη διδακτική των Φυσικών Επιστημών εξετάζει τις εναλλακτικές αντιλήψεις που αναπτύσσουν οι μαθητές για θεμελιώδεις έννοιες. Ορισμένες αντιλήψεις εμφανίζουν ανθεκτικότητα και εσωτερική συνοχή, ενώ άλλες εξαρτώνται από το πλαίσιο της ερώτησης. Επομένως, μια λανθασμένη επιλογή δεν αρκεί από μόνη της για να τεκμηριώσει σταθερή παρανόηση· χρειάζεται εξέταση της αιτιολόγησης και των σχέσεων μεταξύ απαντήσεων.

Οι μελέτες των Driver et al. (1985), Viennot (1979) και Galili (1993) αποτελούν βιβλιογραφικό υπόβαθρο για τη μελέτη των μαθητικών αντιλήψεων και των δυσκολιών στην κατανόηση της μηχανικής. Στην παρούσα εργασία, η καθημερινή ελληνική έκφραση «ζυγίζω 70 κιλά» χρησιμοποιείται ως ενδεικτικό παράδειγμα της γλωσσικής σύγχυσης μάζας και βάρους.

2.2.1 Ταξινόμηση Παρανοήσεων

Για τις ανάγκες της παρούσας ανάλυσης, οι υπό εξέταση παρανοήσεις ομαδοποιούνται ενδεικτικά σε εμπειρικές, γλωσσικές και εννοιολογικές. Η ομαδοποίηση χρησιμοποιείται ως λειτουργικό σχήμα ανάλυσης, με την παραδοχή ότι οι κατηγορίες μπορούν να αλληλεπικαλύπτονται και ότι η προέλευση κάθε μαθητικής αντίληψης απαιτεί ειδικότερη εμπειρική διερεύνηση.

Κατηγορία	Περιγραφή	Παράδειγμα στη Μάζα-Βάρος
Εμπειρικές	Βάσει καθημερινής εμπειρίας	«Στη Σελήνη είμαστε ελαφρύτεροι → η μάζα μειώνεται»
Γλωσσικές	Βάσει καθημερινής γλώσσας	«Ζυγίζω 70 κιλά» → μάζα = βάρος
Εννοιολογικές	Σύγχυση μεταξύ εννοιών	«Βάρος = εγγενής ιδιότητα, δεν αλλάζει»

Πίνακας 2.3: Κατηγορίες Παρανοήσεων

2.2.2 Παρανοήσεις στην Ενότητα Μάζα-Βάρος

Στην ενότητα «Μάζα-Βάρος» επιλέγονται τέσσερις ενδεικτικές παρανοήσεις ως σημεία αναφοράς. Ο Πίνακας 2.4 αποτυπώνει το περιεχόμενό τους και τον αντίστοιχο επιστημονικό διαχωρισμό. Η γενική βιβλιογραφία της Ενότητας 2.2 παρέχει το θεωρητικό υπόβαθρο· δεν αποδίδεται σε κάθε παρανόηση συγκεκριμένος βαθμός ανθεκτικότητας χωρίς ειδική εμπειρική τεκμηρίωση.

Κωδ.	Ενδεικτική παρανόηση	Επιστημονικός διαχωρισμός
M1	«Η μάζα αλλάζει στη Σελήνη»	Η μεταφορά του ίδιου σώματος δεν αλλάζει τη μάζα του· το βάρος εξαρτάται από το βαρυτικό πεδίο.
M2	«Μάζα = βάρος»	Η μάζα και το βάρος είναι διαφορετικά φυσικά μεγέθη, με διαφορετικές μονάδες.
M3	«Το βάρος είναι σταθερή ιδιότητα του σώματος»	Το βάρος του ίδιου σώματος μεταβάλλεται όταν αλλάζει η ένταση του βαρυτικού πεδίου.

Κωδ.	Ενδεικτική παρανόηση	Επιστημονικός διαχωρισμός
M4	«Μεγαλύτερος όγκος σημαίνει πάντα μεγαλύτερη μάζα»	Η σχέση μάζας και όγκου εξαρτάται και από την πυκνότητα του υλικού.

Πίνακας 2.4: Ενδεικτικές Παρανοήσεις στην Ενότητα Μάζα-Βάρος

Το Ph-D A.I. ζητά την αιτιολόγηση της επιλογής A/B/C/D πριν γνωστοποιήσει αν είναι σωστή ή λανθασμένη. Το γλωσσικό μοντέλο αναλύει τον συλλογισμό και μπορεί να επισημάνει ενδείξεις παρανόησης, ακόμη και όταν η τελική επιλογή είναι σωστή. Η επισημάνση αυτή αξιοποιείται ως ερμηνεία της συγκεκριμένης απάντησης βάσει των σημειώσεων και των οδηγιών του συστήματος.

2.2.3 Θεωρία Εννοιολογικής Αλλαγής

Στο κλασικό μοντέλο των Posner et al. (1982), η εννοιολογική αλλαγή αφορά την αναδιοργάνωση των αρχικών αντιλήψεων όταν αυτές δεν επαρκούν για την ερμηνεία ενός φαινομένου. Το μοντέλο προτείνει τέσσερις συνθήκες για την αποδοχή μιας νέας αντίληψης:

1. Δυσaréσκεια με την παλιά αντίληψη
2. Κατανοητότητα της νέας έννοιας
3. Πειστικότητα της νέας έννοιας
4. Γονιμότητα της νέας έννοιας

Ο σχεδιασμός του Ph-D A.I. επιδιώκει να δημιουργεί ευκαιρίες εξέτασης των αρχικών αντιλήψεων. Η αντιπαραβολή της αιτιολόγησης με τις σχολικές σημειώσεις μπορεί να αναδείξει περιορισμούς της αρχικής εξήγησης, ενώ η αναλυτική παρουσίαση και τα παραδείγματα στοχεύουν στην κατανοητότητα και την πειστικότητα της επιστημονικής έννοιας. Οι υποθετικές περιπτώσεις δείχνουν πώς η νέα έννοια εφαρμόζεται και σε άλλες καταστάσεις. Το μοντέλο των Posner et al. (1982) χρησιμοποιείται εδώ ως οδηγός για τον σχεδιασμό του συστήματος.

2.3 Τεχνητή Νοημοσύνη ως Παιδαγωγικό Εργαλείο

Η τεχνητή νοημοσύνη στην εκπαίδευση (AIED — Artificial Intelligence in Education) περιλαμβάνει διαφορετικές προσεγγίσεις για τη μοντελοποίηση της γνώσης, την προσαρμογή δραστηριοτήτων και την παροχή ανατροφοδότησης. Τα ευφυή συστήματα διδασκαλίας και

τα συστήματα μεγάλων γλωσσικών μοντέλων συνιστούν διαφορετικές, ενίοτε συμπληρωματικές, τεχνολογικές κατευθύνσεις.

2.3.1 Από τα Παραδοσιακά ITS στα Σύγχρονα LLM

Τα Intelligent Tutoring Systems (ITS, ευφυή συστήματα διδασκαλίας) συνήθως αξιοποιούν ρητές αναπαραστάσεις του γνωστικού πεδίου, του μαθητή και των διδακτικών ενεργειών. Η προσαρμογή μπορεί να βασίζεται σε κανόνες ή σε υπολογιστικά μοντέλα. Οι δυνατότητες επεξεργασίας φυσικής γλώσσας διαφέρουν ανά σύστημα: το AutoTutor, για παράδειγμα, υποστηρίζει διάλογο φυσικής γλώσσας και αλληλεπίδραση μικτής πρωτοβουλίας (Graesser et al., 2005). Συνεπώς, δεν είναι ακριβής η γενική ταύτιση όλων των ITS με στατικά μηνύματα ή περιορισμό σε λέξεις-κλειδιά.

Τα μεγάλα γλωσσικά μοντέλα (Large Language Models — LLMs) προσφέρουν ευέλικτη παραγωγή φυσικού λόγου και δυνατότητες διαλογικής ανατροφοδότησης. Η προσαρμογή τους εξαρτάται από τις οδηγίες, το διαθέσιμο ιστορικό και το εκπαιδευτικό υλικό που τους παρέχεται. Παραμένουν προκλήσεις η παραγωγή ανακριβών πληροφοριών, η διαφάνεια των κρίσεων και η αξιοπιστία της αξιολόγησης (Kasneci et al., 2023). Ένα δομημένο prompt μπορεί να καθοδηγεί συστηματικά την ανάλυση, ενώ η ορθότητα και η σταθερότητα των αποτελεσμάτων χρειάζονται πρόσθετο έλεγχο.

Ο Πίνακας 2.5 συγκρίνει ενδεικτικές αρχιτεκτονικές τάσεις των ITS και των συστημάτων LLM. Οι περιγραφές αποτυπώνουν χαρακτηριστικές τάσεις των δύο κατηγοριών, αναγνωρίζοντας ότι οι επιμέρους υλοποιήσεις μπορεί να διαφοροποιούνται σημαντικά.

Χαρακτηριστικό	ITS με ρητά μοντέλα	Συστήματα με LLM
Φυσική γλώσσα	Διαφέρει ανά σύστημα· υπάρχουν διαλογικά ITS όπως το AutoTutor.	Ευέλικτη παραγωγή και επεξεργασία φυσικού λόγου.
Προσαρμογή	Μοντέλα μαθητή και διδακτικές πολιτικές ή κανόνες.	Οδηγίες, διαθέσιμο ιστορικό και πρόσθετοι μηχανισμοί της εφαρμογής.
Ανατροφοδότηση	Μπορεί να είναι προκαθορισμένη ή να παράγεται από μοντέλα του πεδίου.	Παραγόμενες εξηγήσεις με ανάγκη ελέγχου της ορθότητάς τους.

Χαρακτηριστικό	ITS με ρητά μοντέλα	Συστήματα με LLM
Συλλογισμός	Μοντελοποίηση βημάτων ή διάλογος, ανάλογα με τον σχεδιασμό.	Ανάλυση αιτιολόγησης μέσω οδηγιών, χωρίς εγγυημένη αξιοπιστία.
Έλεγχος λειτουργίας	Ρητοί κανόνες και αναπαραστάσεις του πεδίου.	Prompt και κώδικας.

Πίνακας 2.5: Σύγκριση ITS με Ρητά Μοντέλα και Συστημάτων με LLM

Το Ph-D A.I. συνδυάζει την παροχή σχολικών σημειώσεων στο γλωσσικό μοντέλο με τη δομημένη ανάλυση του μαθητικού συλλογισμού. Ο όρος RAG χρησιμοποιείται εδώ με την έννοια μιας απλοποιημένης, τοπικής προσέγγισης ανάκτησης του αρχείου σημειώσεων του επιλεγμένου κεφαλαίου και ένταξής του στο context του μοντέλου (το κείμενο που λαμβάνει το μοντέλο σε κάθε αίτημα). Η υλοποίηση δεν περιλαμβάνει διανυσματική αναζήτηση: ο κώδικας παρέχει έως 20.000 χαρακτήρες του αρχείου και το prompt ζητά οι απαντήσεις να στηρίζονται αποκλειστικά στο διαθέσιμο υλικό. Η χρήση των σημειώσεων στηρίζεται στις οδηγίες του prompt. Παράλληλα, το VCLRCI-Q οργανώνει την περιγραφική αξιολόγηση της αιτιολόγησης σε πέντε διαστάσεις.

Ο Πίνακας 2.6 παρουσιάζει διαφορετικές προσεγγίσεις εκπαιδευτικής υποστήριξης, χωρίς συγκριτική αξιολόγηση της αποτελεσματικότητάς τους. Η περιγραφή του AutoTutor βασίζεται στους Graesser et al. (2005), ενώ οι περιγραφές του Khanmigo (Khan Academy, n.d.) και του ASSISTments (ASSISTments, n.d.) βασίζονται στις επίσημες σελίδες τους. Η περιγραφή του Ph-D A.I. βασίζεται στο system prompt και στον κώδικα της παρούσας εργασίας.

Σύστημα	Προσέγγιση	Υποστήριξη ανατροφοδότησης	Μορφή αλληλεπίδρασης
AutoTutor	Διαλογικό ITS	Εξέταση απαντήσεων στο πλαίσιο διδακτικού διαλόγου.	Φυσική γλώσσα και μικτή πρωτοβουλία.
Khanmigo	Εκπαιδευτικός βοηθός με τεχνητή νοημοσύνη	Ερωτήσεις και καθοδήγηση για διερεύνηση της λύσης.	Διαλογική υποστήριξη του μαθητή.

Σύστημα	Προσέγγιση	Υποστήριξη ανατροφοδότησης	Μορφή αλληλεπίδρασης
ASSISTments	Πλατφόρμα εξάσκησης	Άμεση ανατροφοδότηση, υποδείξεις και αναφορές στον εκπαιδευτικό.	Απαντήσεις σε δραστηριότητες και βοήθεια κατά την εξάσκηση.
Ph-D A.I.	LLM με σημειώσεις και δομημένο prompt	Ανάλυση αιτιολόγησης και δείκτες VCLRCI-Q στα συνθετικά σενάρια.	Τρεις ρόλοι του ίδιου μοντέλου και περιορισμένη ακολουθία διαλόγου.

Πίνακας 2.6: Ενδεικτικές Προσεγγίσεις Εκπαιδευτικών Συστημάτων

Η συμβολή της παρούσας εργασίας εστιάζει στον συγκεκριμένο συνδυασμό σχολικών σημειώσεων, τριών παιδαγωγικών ρόλων και δομημένων κριτηρίων αξιολόγησης σε ένα πρωτότυπο εκπαιδευτικό σύστημα για τη διδασκαλία της Φυσικής. Η λειτουργία του διερευνάται σε συνθετικά σενάρια, τα οποία αποτελούν το αρχικό στάδιο αξιολόγησης του συστήματος.

2.4 Ποιότητα Ανατροφοδότησης και Μάθηση

Η ανατροφοδότηση είναι κεντρικό στοιχείο κάθε αποτελεσματικής μαθησιακής διαδικασίας, ιδίως όταν στοχεύει στην εννοιολογική αλλαγή σε απαιτητικά πεδία όπως η Φυσική.

2.4.1 Το Μοντέλο Hattie & Timperley (2007)

Οι Hattie και Timperley (2007), στο *The Power of Feedback*, διακρίνουν τέσσερα επίπεδα ανατροφοδότησης: το έργο ή την απάντηση (task), τη διαδικασία επίλυσης (process), την αυτορρύθμιση (self-regulation) και το πρόσωπο του μαθητή (self). Η ανατροφοδότηση στο έργο αφορά το περιεχόμενο της απάντησης, ενώ εκείνη στη διαδικασία αφορά τις στρατηγικές και τις σχέσεις που στηρίζουν την επίλυση. Η αυτορρύθμιση συνδέεται με την παρακολούθηση και την αξιολόγηση της μαθησιακής πορείας.

Το Ph-D A.I. σχεδιάστηκε κυρίως για ανατροφοδότηση στο έργο και στη διαδικασία: γνωστοποιεί την ορθότητα της επιλογής μετά την αιτιολόγηση και εξηγεί τον συλλογισμό με βάση τις σημειώσεις. Οι διερευνητικές ερωτήσεις των ρόλων προσφέρουν ευκαιρίες αναστοχασμού, ενώ η μελέτη εστιάζει στα επίπεδα του έργου και της διαδικασίας.

Η εξέταση της αιτιολόγησης διευρύνει την ανατροφοδότηση πέρα από την ορθότητα της τελικής επιλογής και επιτρέπει την περιγραφή εννοιολογικών δυσκολιών και λογικών βημάτων σε κάθε απάντηση.

2.4.2 Επεξηγηματική Ανατροφοδότηση (Elaborated Feedback)

Η Shute (2008) εξετάζει τη διαμορφωτική ανατροφοδότηση ως πληροφορία που μπορεί να υποστηρίξει τη βελτίωση της σκέψης και της επίδοσης. Η επεξηγηματική ανατροφοδότηση περιλαμβάνει στοιχεία πέρα από τη γνωστοποίηση σωστού ή λάθους, όπως εξηγήσεις και υποδείξεις. Στον σχεδιασμό του Ph-D A.I. αξιοποιούνται οι ακόλουθες αρχές, τις οποίες η παρούσα εργασία συνθέτει με βάση τη βιβλιογραφία για την επεξηγηματική ανατροφοδότηση:

1. **Εξειδίκευση:** αναφορά στο συγκεκριμένο σημείο της αιτιολόγησης που χρειάζεται εξέταση.
2. **Επεξήγηση:** παρουσίαση του επιστημονικού συλλογισμού και του λόγου για τον οποίο στηρίζει τη σωστή απάντηση.
3. **Παραδείγματα:** σύνδεση της έννοιας με κατάλληλες σχολικές ή καθημερινές περιπτώσεις.
4. **Κατάλληλος χρονισμός:** παροχή ανατροφοδότησης αφού ο μαθητής έχει εκφράσει την αιτιολόγησή του.

Σύμφωνα με το system prompt, ο Κύριος Γιώργος δίνει σύντομη εξήγηση με παράδειγμα και αναφέρεται στον συλλογισμό του μαθητή. Η ανατροφοδότηση ακολουθεί την αιτιολόγηση και συνοδεύεται από τον πίνακα με τους δείκτες V, L, R και Clarity. Η εφαρμογή αυτών των αρχών στα συνθετικά σενάρια εξετάζεται στο Κεφάλαιο 5.

2.4.3 Χρονισμός και Προσαρμογή

Η Shute (2008) επισημαίνει ότι ο κατάλληλος χρονισμός της ανατροφοδότησης εξαρτάται από το έργο και τα χαρακτηριστικά του μαθητή. Στο Ph-D A.I. η ανατροφοδότηση παρέχεται μετά την επιλογή και την αιτιολόγηση κάθε απάντησης, ενώ στο τέλος παρουσιάζεται συνολική αποτίμηση. Η τελική σύνοψη λειτουργεί ως συγκεντρωτική παρουσίαση της συνεδρίας. Η διατύπωση της εξήγησης μπορεί να προσαρμόζεται στην αιτιολόγηση· η δυσκολία των ερωτήσεων επιλέγεται κατά τη δημιουργία του κουίζ, με βάση τις οδηγίες και το διαθέσιμο ιστορικό προηγούμενων συνεδριών.

2.5 Το Πλαίσιο Αξιολόγησης VCLRCI-Q — Θεωρητική Θεμελίωση

Το VCLRCI-Q αποτελεί πλαίσιο που διαμορφώθηκε για την παρούσα εφαρμογή, με βιβλιογραφικές αφετηρίες στη διδακτική της Φυσικής, την επιχειρηματολογία και την ανατροφοδότηση. Οι λειτουργικοί ορισμοί και οι κανόνες βαθμολόγησης καθορίζονται στο system prompt, όπου το πλαίσιο υλοποιείται ως ρουμπρικά αξιολόγησης· γι' αυτό στην εργασία χρησιμοποιούνται και οι δύο όροι. Η βιβλιογραφία παρέχει τον θεωρητικό προσανατολισμό του πλαισίου, ενώ η ψυχομετρική επικύρωση των δεικτών αποτελεί αντικείμενο μελλοντικής έρευνας.

Ο δείκτης Validity (V, εγκυρότητα) εξετάζει την επιστημονική ορθότητα της αιτιολόγησης, την ορθή χρήση των εννοιών και τη συμφωνία με τις σχολικές σημειώσεις. Συνδέεται θεματικά με τη μελέτη των παρανοήσεων (Driver et al., 1985· Galili, 1993). Ο Consistency (C, συνέπεια) εξετάζει στο τέλος του κουίζ αν οι ισχυρισμοί των απαντήσεων είναι συμβατοί μεταξύ τους. Ένα σταθερό αλλά λανθασμένο εννοιολογικό μοντέλο μπορεί να έχει υψηλή συνέπεια, ενώ η εναλλαγή σωστών και λανθασμένων επιλογών δεν αρκεί από μόνη της για να δείξει αντίφαση. Επίσης, η αιτιολογημένη αναθεώρηση μιας ιδέας διακρίνεται από την ασυνεπή χρήση της. Ο Logic (L, λογική) εξετάζει την αλληλουχία των βημάτων, τη λογική σύνδεση και την εσωτερική συνοχή κάθε αιτιολόγησης, σε συνάφεια με την επιχειρηματολογία των Chin και Osborne (2010).

Ο Regulation (R, εκπαιδευτική καταλληλότητα), όπως ορίζεται στο prompt, εξετάζει αν ο συλλογισμός είναι κατάλληλος για τη δηλωμένη τάξη, την ύλη και τη σχολική μεθοδολογία, χωρίς αδικαιολόγητα προχωρημένη ή υπερβολικά επιφανειακή προσέγγιση. Ο δείκτης αφορά την εκπαιδευτική καταλληλότητα της απάντησης και όχι τη μεταγνωστική επίγνωση ή την αυτορρύθμιση του μαθητή. Στη διεπαφή του συστήματος εμφανίζεται ως «R (Επίπεδο)»· ο όρος Regulation αναφέρεται εδώ στη ρύθμιση του συλλογισμού ως προς το επίπεδο της βαθμίδας. Ο Clarity (Cl, σαφήνεια) εξετάζει τη δομή, την κατανοητότητα, την ορολογία, την απουσία αμφισημίας και τη σύνδεση της διατύπωσης με το ερώτημα. Η σαφήνεια διακρίνεται από την επιστημονική ορθότητα, σε συνάφεια με τη σημασία της κατανοητής πληροφορίας στην ανατροφοδότηση (Shute, 2008).

Η διάκριση των δεικτών επιτρέπει να περιγραφεί ένας συλλογισμός που είναι σαφής και συνεπής αλλά επιστημονικά λανθασμένος. Σύμφωνα με το prompt, οι V, L, R και Cl προκύπτουν από 20 ελέγχους ανά αιτιολόγηση, πέντε ανά δείκτη, και ο C από πέντε επιπλέον ελέγχους στο τέλος του κουίζ. Οι επιμέρους έλεγχοι δεν αποθηκεύονται ως χωριστές

αριθμητικές τιμές στο Q_RECORD (τη δομημένη εγγραφή αξιολόγησης κάθε ερώτησης). Το μοντέλο δίνει απευθείας τις τιμές V, L, R και CI και εξηγεί με λόγια κάθε βαθμό στο score_analysis (τη λεκτική αιτιολόγηση των βαθμών), αναφέροντας τους ελέγχους που δεν ικανοποιήθηκαν πλήρως. Από αυτές τις τιμές και τον C προκύπτει ο συνολικός δείκτης Q κάθε ερώτησης. Όλες τις κρίσεις τις κάνει το ίδιο γλωσσικό μοντέλο και αφορούν τον συλλογισμό του μαθητή. Στην παρούσα εργασία οι τιμές αξιοποιούνται περιγραφικά, στο πλαίσιο της λειτουργικής διερεύνησης του συστήματος.

2.6 Σύνοψη Θεωρητικού Πλαισίου

Το θεωρητικό πλαίσιο συνδέει τη διαλογική διδασκαλία (Alexander, 2008· Chin & Osborne, 2010), τη μελέτη των παρανοήσεων και της εννοιολογικής αλλαγής στη Φυσική (Driver et al., 1985· Viennot, 1979· Galili, 1993· Posner et al., 1982) και την εκπαιδευτική αξιοποίηση της τεχνητής νοημοσύνης και της ανατροφοδότησης (Luckin et al., 2016· Kasneci et al., 2023· Hattie & Timperley, 2007· Shute, 2008). Οι αρχές αυτές καθοδηγούν τον σχεδιασμό του Ph-D A.I. και τη λειτουργική διαμόρφωση του VCLRCI-Q. Το επόμενο κεφάλαιο παρουσιάζει την τεχνική υλοποίηση, διακρίνοντας τις οδηγίες του prompt από τους μηχανισμούς του κώδικα.

3. Το Σύστημα Ph-D A.I. (PhysicsDialog AI)

Το κεφάλαιο περιγράφει την τεχνική υλοποίηση του Ph-D A.I. (PhysicsDialog AI). Ο σχεδιασμός αξιοποιεί τις παιδαγωγικές αρχές του Κεφαλαίου 2, με έμφαση στην αιτιολόγηση, στη διαλογική αλληλεπίδραση και στην επεξηγηματική ανατροφοδότηση. Η περιγραφή βασίζεται στον κώδικα HTML/JavaScript και PHP και στο system prompt PHYSICS-QUIZ RAG Edition v2.0 (τις βασικές οδηγίες προς το μοντέλο), όπως ήταν κατά την εκτέλεση των σεναρίων του Κεφαλαίου 4, και διακρίνει όσα εκτελεί ο κώδικας από όσα ζητείται να εφαρμόσει το γλωσσικό μοντέλο.

3.1 Αρχιτεκτονική Συστήματος

Το Ph-D A.I. υλοποιείται ως διαδικτυακή εφαρμογή με frontend (διεπαφή χρήστη) σε HTML/CSS/JavaScript και backend (διακομιστή εφαρμογής) σε PHP. Ο φυλλομετρητής διατηρεί τα μηνύματα της ενεργής συνομιλίας και στέλνει το ιστορικό σε κάθε αίτημα. Το PHP φορτώνει τις οδηγίες και τις σημειώσεις, επικοινωνεί με το API (τη διεπαφή προγραμματισμού του παρόχου του μοντέλου) και αποθηκεύει τις τελικές συνόψεις σε αρχεία JSON (μορφή δομημένων δεδομένων).

Στρώμα	Τεχνολογία	Ρόλος
Frontend (Διεπαφή Χρήστη)	HTML5, CSS3, JavaScript	Διεπαφή χρήστη, απεικόνιση μαθηματικών συμβόλων (MathJax) και τοπική αποθήκευση δεδομένων.
Backend (Εξυπηρετητής Εφαρμογής)	PHP	Διαχείριση αιτημάτων, επικοινωνία με το γλωσσικό μοντέλο και διασύνδεση με το γνωστικό υλικό.
AI Engine (Μηχανή Τεχνητής Νοημοσύνης)	Claude Haiku 4.5	Δημιουργία ερωτήσεων, αξιολόγηση απαντήσεων, παροχή ανατροφοδότησης και διαχείριση διαλόγου.
Data Layer (Στρώμα Δεδομένων)	Αρχεία JSON	Μόνιμη αποθήκευση τελικών συνόψεων ανά username (όνομα χρήστη). τα αναλυτικά Q_RECORD (εγγραφές αξιολόγησης ανά ερώτηση) διατηρούνται στον φυλλομετρητή.

Στρώμα	Τεχνολογία	Ρόλος
Knowledge Layer (Στρώμα Γνώσης)	Αρχεία TXT	Άμεση ανάκτηση αρχείου κεφαλαίου και παροχή έως 20.000 χαρακτήρων στο μοντέλο.

Πίνακας 3.1: Αρχιτεκτονικά Στρώματα του Ph-D A.I. (PhysicsDialog AI)

3.1.1 Frontend — Διεπαφή Χρήστη

Η διεπαφή χρήστη αναπτύχθηκε με τεχνολογίες HTML5, CSS3 και JavaScript και αποτελεί το κύριο μέσο αλληλεπίδρασης του μαθητή με το σύστημα. Υποστηρίζει την εισαγωγή στοιχείων χρήστη και την επιλογή διδακτικής ενότητας, ενώ παρέχει περιβάλλον διαλογικής επικοινωνίας με εμφάνιση μαθηματικών συμβολισμών μέσω LaTeX (γλώσσας γραφής μαθηματικών τύπων) και μορφοποιημένου εκπαιδευτικού περιεχομένου.

Η συνομιλία διατηρείται σε μεταβλητή JavaScript. Από κάθε απάντηση του μοντέλου εξάγονται τα Q_RECORD, οι έγκυρες εγγραφές αποθηκεύονται στο localStorage (τον τοπικό χώρο αποθήκευσης του φυλλομετρητή) με κλειδί ανά αναγνωριστικό συνεδρίας και τα αντίστοιχα τμήματα αφαιρούνται πριν από την εμφάνιση. Ο μαθητής δεν βλέπει αυτά τα δεδομένα σε μορφή JSON, βλέπει όμως τους πίνακες βαθμολογίας και την τελική αξιολόγηση. Παρέχονται επίσης γρήγορες επιλογές συνέχισης ή επανεξήγησης και ένδειξη αναμονής κατά το αίτημα.

Με τον τρόπο αυτό, η διεπαφή δεν λειτουργεί μόνο ως μέσο παρουσίασης περιεχομένου, αλλά και ως ενδιάμεσο επίπεδο διαχείρισης της αλληλεπίδρασης και των δεδομένων που παράγονται κατά τη μαθησιακή διαδικασία.

3.1.2 Διακομιστής Εφαρμογής (Backend – PHP)

Ο διακομιστής physics_quiz.php διαβάζει από το εισερχόμενο JSON τα messages (μηνύματα της συνομιλίας), username, book (βιβλίο) και chapter (κεφάλαιο). Το frontend στέλνει επιπλέον session_id (αναγνωριστικό συνεδρίας) και questions (ερωτήσεις), που προορίζονται για μελλοντική χρήση. Ο διακομιστής φορτώνει το system_prompt_quiz.txt, το αρχείο σημειώσεων που αντιστοιχεί στο βιβλίο και το κεφάλαιο, καθώς και έως τρεις πρόσφατες αποθηκευμένες συνόψεις του ίδιου username.

Με αυτά τα δεδομένα κατασκευάζει το context (το σύνολο του κειμένου που στέλνεται στο μοντέλο) και στέλνει ένα αίτημα στο Anthropic Messages API με το μοντέλο claude-haiku-4-5-20251001 και max_tokens=8000 (μέγιστο μήκος απάντησης). Μετά την απάντηση ενώνει

τα παραγόμενα τμήματα κειμένου, αναζητά το QUIZ_SUMMARY και επιχειρεί να το αναλύσει ως JSON.

Εφόσον εντοπιστεί αναγνώσιμη σύνοψη, το PHP δημιουργεί νέο αναγνωριστικό και χρονική σήμανση για τη μόνιμη εγγραφή και επιχειρεί να την προσθέσει στο ιστορικό του username. Επιστρέφει reply (την απάντηση), username, saved (ένδειξη αποθήκευσης) και usage (στοιχεία κατανάλωσης).

3.1.3 Γλωσσικό Μοντέλο (Claude Haiku 4.5)

Στον κώδικα της εφαρμογής έχει οριστεί το Claude Haiku 4.5 με αναγνωριστικό claude-haiku-4-5-20251001. Η επιλογή του ταιριάζει σε μια εφαρμογή με συχνή διαλογική αλληλεπίδραση· η σύγκριση με άλλα μοντέλα δεν αποτέλεσε αντικείμενο της παρούσας μελέτης.

Το ίδιο μοντέλο παράγει τις ερωτήσεις, αναλύει την αιτιολόγηση, αποδίδει βαθμούς και συντάσσει τις παρεμβάσεις των τριών ρόλων. Η ποιότητα αυτών των λειτουργιών εξετάζεται περιγραφικά στα συνθετικά σενάρια του Κεφαλαίου 5.

Το όριο max_tokens=8000 αφορά το μέγιστο πλήθος tokens (μονάδων κειμένου) της παραγόμενης απάντησης και όχι το μέγεθος του εισερχόμενου context. Η απάντηση μπορεί να περιλαμβάνει μαθηματική σημειογραφία, εξηγήσεις και δομημένα σήματα.

Κάθε κλήση στο API είναι ανεξάρτητη και δεν χρησιμοποιείται extended thinking (εκτεταμένη συλλογιστική πριν από την απάντηση). Επομένως, όσα το prompt ορίζει ως εσωτερικά, δηλαδή η γνωστική βάση (knowledge_base), το σύνολο ερωτήσεων, η ανάλυση του συλλογισμού σε βήματα (Reasoning Graph), οι επιμέρους έλεγχοι και το αρχείο συνέπειας (consistency_log), δεν διατηρούνται μεταξύ κλήσεων ούτε καταγράφονται. Σε κάθε απάντηση το μοντέλο βασίζεται μόνο στο system prompt και στο ιστορικό της συνομιλίας που του στέλνει η διεπαφή. Το ιστορικό περιλαμβάνει ολόκληρες τις προηγούμενες απαντήσεις του μοντέλου, μαζί με τα Q_RECORD που δεν εμφανίζονται στον μαθητή. Οι επιμέρους έλεγχοι λειτουργούν ως κριτήρια της κρίσης του μοντέλου και όχι ως χωριστά υπολογισμένες τιμές.

3.2 RAG (Retrieval-Augmented Generation) — Ανάκτηση-Ενισχυμένη Παραγωγή

Στην παρούσα εφαρμογή ο όρος RAG χρησιμοποιείται για την ανάκτηση του αρχείου σημειώσεων του επιλεγμένου κεφαλαίου και την ένταξή του στο context του μοντέλου. Η ανάκτηση είναι άμεση, βάσει ονόματος αρχείου. Δεν περιλαμβάνει embeddings (διανυσματικές αναπαραστάσεις κειμένου), διανυσματική βάση ή σημασιολογική αναζήτηση αποσπασμάτων. Οι σημειώσεις αποτελούν τη σχολική γνωστική βάση της εφαρμογής.

3.2.1 Δομή Αρχείων RAG

Τα αρχεία γνώσης βρίσκονται στον κατάλογο `/api/notes_txt/` και ακολουθούν το σχήμα `{book}_{chapter}.txt`, όπου book είναι το βιβλίο και chapter ο αριθμός του κεφαλαίου που επιλέγει ο μαθητής στη διεπαφή. Έτσι το αίτημα αντιστοιχίζεται σε συγκεκριμένο αρχείο, χωρίς αναζήτηση με βάση το κείμενο της ερώτησης.

Σε κάθε αίτημα ο διακομιστής φορτώνει πρώτα το αντίστοιχο αρχείο σημειώσεων και στη συνέχεια καλεί το μοντέλο. Αν το αρχείο δεν υπάρχει, καταγράφει το συμβάν στο error log (αρχείο σφαλμάτων) και ενημερώνει το μοντέλο ότι οι σημειώσεις δεν βρέθηκαν. Στην περίπτωση αυτή το system prompt ορίζει ότι το κουίζ δεν ξεκινά: το μοντέλο ενημερώνει τον μαθητή ότι οι σημειώσεις είναι ελλιπείς και του ζητά να επιλέξει άλλο κεφάλαιο.

Η προσθήκη υλικού για βιβλία και κεφάλαια που υποστηρίζονται ήδη γίνεται με νέα αρχεία TXT. Για νέα βιβλία ή κεφάλαια απαιτείται επιπλέον η προσθήκη τους στη λίστα επιλογών της διεπαφής. Σε κάθε περίπτωση, το περιεχόμενο και το μέγεθος των αρχείων πρέπει να ελέγχονται, ώστε το υλικό που φτάνει στο μοντέλο να καλύπτει τη διδακτική ενότητα.

3.2.2 Ενσωμάτωση του Γνωστικού Υλικού στο Prompt

Το PHP ενσωματώνει έως τους πρώτους 20.000 χαρακτήρες του αρχείου με `mb_substr($txt_content, 0, 20000)`. Αν το αρχείο είναι μεγαλύτερο, το υπόλοιπο δεν παρέχεται στο μοντέλο και η αποκοπή μπορεί να γίνει στη μέση μιας πρότασης. Το block (τμήμα κειμένου) συνοδεύεται από οδηγίες για αποκλειστική χρήση των σημειώσεων και αποφυγή web search (αναζήτησης στο διαδίκτυο). Η τήρηση του περιορισμού αυτού στηρίζεται στις οδηγίες του prompt.

Το system prompt ζητά από το μοντέλο να εξαγάγει εσωτερικά ορισμούς, σχέσεις και βασικές έννοιες από τις σημειώσεις. Αν δεν υπάρχουν, το κουίζ δεν ξεκινά και ζητείται από τον μαθητή να επιλέξει άλλο κεφάλαιο. Όταν οι σημειώσεις δεν περιέχουν τύπους,

προβλέπονται μόνο εννοιολογικές ερωτήσεις. Κατά τη δημιουργία κάθε ερώτησης ζητείται αυτοέλεγχος της λύσης, της σωστής επιλογής και όλων των distractors (λανθασμένων επιλογών). Αν ο έλεγχος αποτύχει, προβλέπονται έως τρεις προσπάθειες και στη συνέχεια παράλειψη της ερώτησης, με μείωση του αριθμού ερωτήσεων.

Οι έλεγχοι αυτοί εκτελούνται από το ίδιο το μοντέλο, σύμφωνα με τις οδηγίες του prompt, και αποσκοπούν στον περιορισμό των σφαλμάτων.

3.2.3 Σχεδιαστικές δυνατότητες της παροχής σημειώσεων

Η χρήση επιλεγμένου εκπαιδευτικού υλικού προσφέρει κοινό σημείο αναφοράς για ερωτήσεις και ανατροφοδότηση. Στον Πίνακα 3.2 συγκρίνονται οι σχεδιαστικές δυνατότητες της συγκεκριμένης υλοποίησης με μια κλήση LLM χωρίς παρεχόμενες σημειώσεις.

Παράμετρος	Χωρίς παρεχόμενες σημειώσεις	Με σημειώσεις στο Ph-D A.I.
Γνωστική αναφορά	Γενική γνώση μοντέλου και οδηγίες.	Επιλεγμένο κείμενο κεφαλαίου και οδηγία αποκλειστικής χρήσης του.
Σχολικό περιεχόμενο	Η συνάφεια εξαρτάται από το prompt και το μοντέλο.	Υπάρχει συγκεκριμένη είσοδος σχολικού υλικού, με όριο 20.000 χαρακτήρων.
Ακρίβεια	Χρειάζεται έλεγχο των εξόδων.	Χρειάζεται επίσης έλεγχο· η μείωση σφαλμάτων δεν έχει μετρηθεί συγκριτικά.
Εκπαιδευτικό επίπεδο	Καθορίζεται με οδηγίες.	Καθορίζεται από τις οδηγίες για την τάξη και από το υλικό του κεφαλαίου.
Έλεγχος εισόδου	Έλεγχος του prompt.	Ο σχεδιαστής ελέγχει το prompt και το αρχείο σημειώσεων που λαμβάνει το μοντέλο.

Πίνακας 3.2: Παροχή Σημειώσεων και Γενική Γνώση LLM

3.3 Prompt Caching (Προσωρινή Αποθήκευση Οδηγιών)

Ο κώδικας χρησιμοποιεί `cache_control` με `type="ephemeral"` (σήμανση προσωρινής αποθήκευσης) σε τέσσερα σημεία: στο system prompt, στο block σημειώσεων όταν υπάρχει, στο block ιστορικού και στο τελευταίο προηγούμενο μήνυμα assistant (του μοντέλου). Η

σήμανση επιτρέπει στον πάροχο να επαναχρησιμοποιεί το κοινό αρχικό τμήμα του prompt από προηγούμενα αιτήματα.

Το πρώτο block περιέχει το system prompt, με τους παιδαγωγικούς ρόλους, τη ροή του κουίζ, τη ρουμπρίκα και τα σήματα καταγραφής. Το δεύτερο περιέχει τις σημειώσεις του κεφαλαίου, έως 20.000 χαρακτήρες, με τις συνοδευτικές οδηγίες. Το block ιστορικού περιέχει έως τρεις πρόσφατες συνόψεις του username, με θέμα, τάξη και συνολικούς δείκτες· αν δεν υπάρχει ιστορικό, δηλώνεται προεπιλεγμένο μέτριο επίπεδο. Οι συνόψεις δεν φιλτράρονται ως προς το τρέχον θέμα ή την τάξη.

Το PHP καταγράφει τα στοιχεία χρήσης της cache (προσωρινής μνήμης) που επιστρέφει το API (cache_creation_input_tokens, cache_read_input_tokens, input_tokens) και τα επιστρέφει και στο frontend. Το πλήρες payload αποστέλλεται σε κάθε κλήση, το caching σημαίνει ότι ο πάροχος επαναχρησιμοποιεί τμήματα που έχει ήδη επεξεργαστεί. Η μείωση κόστους και καθυστέρησης εξαρτάται από τις επιτυχείς αναγνώσεις της cache (Anthropic, n.d.).

3.4 Προσαρμοστική Δυσκολία (Adaptive Difficulty)

Η προσαρμογή της δυσκολίας προβλέπεται στο system prompt και βασίζεται στις συνόψεις προηγούμενων συνεδριών. Το PHP παρέχει χωριστά τις τιμές avg_Q (μέσος Q κάθε συνεδρίας) έως τριών συνόψεων· τον μέσο όρο τους και την επιλογή του επιπέδου τα κάνει το μοντέλο, σύμφωνα με τις οδηγίες.

Το prompt ορίζει τρία επίπεδα με βάση τον μέσο avg_Q:

- **κάτω από 0,50:** χαμηλό επίπεδο, με απλές ερωτήσεις και ευκολότερα distractors·
- **από 0,50 έως 0,74:** μέτριο επίπεδο, που αποτελεί και την προεπιλογή όταν δεν υπάρχει ιστορικό·
- **0,75 και πάνω:** υψηλό επίπεδο, με σύνθετες ή πολυβηματικές ερωτήσεις και δυσκολότερα distractors.

Σε κάθε επίπεδο, η διατύπωση και το περιεχόμενο των ερωτήσεων περιορίζονται από τις σημειώσεις και τη δηλωμένη τάξη. Επειδή το ιστορικό είναι ίδιο σε όλη τη διάρκεια ενός κουίζ, η δυσκολία μένει σταθερή μέσα στη συνεδρία και δεν αλλάζει μετά από κάθε

απάντηση. Οι ερωτήσεις παρουσιάζονται μία προς μία. Η δυσκολία επιλέγεται χωρίς να ανακοινώνεται στον μαθητή.

3.5 Διαλογική Δομή — Οι Τρεις Ρόλοι

Το system prompt ορίζει τρεις παιδαγωγικούς ρόλους με διαφορετικές λειτουργίες, ως σχεδιαστική αξιοποίηση της διαλογικής προσέγγισης (Alexander, 2008). Όλες οι παρεμβάσεις παράγονται από το ίδιο μοντέλο. Οι ρόλοι είναι οπτικές μέσα σε μια προκαθορισμένη ακολουθία και όχι ανεξάρτητοι πράκτορες (agents) ή ξεχωριστά μοντέλα.

Μετά την αιτιολόγηση, ο Κύριος Γιώργος γνωστοποιεί την ορθότητα της επιλογής και δίνει εξήγηση 3–5 προτάσεων, με παράδειγμα και σύνδεση με τις σημειώσεις. Ακολουθούν ο ορατός πίνακας V, L, R και Clarity, το μη εμφανές Q_RECORD και ένας γύρος διαλόγου: Ζωή → Κύριος Γιώργος → Ελισάβετ → Κύριος Γιώργος. Στη συνέχεια τίθεται ερώτηση κατανόησης και επιλέγεται συνέχιση ή επανεξήγηση.

Ρόλος	Χαρακτήρας	Παιδαγωγική Λειτουργία	Κανόνες
Κύριος Γιώργος	Κύριος διδάσκων	Παρέχει την κύρια διδασκαλία, εξηγεί τη σωστή απάντηση, αναλύει τα λάθη και συνδέει τη συζήτηση με το εκπαιδευτικό υλικό.	Εξήγηση 3–5 προτάσεων μετά την αιτιολόγηση· σύντομες απαντήσεις στις δύο παρεμβάσεις.
Ζωή	Περίεργη μαθήτρια	Διατυπώνει διερευνητικές ερωτήσεις που επεκτείνουν τη συζήτηση και ενθαρρύνουν βαθύτερη κατανόηση των εννοιών.	Μία ερώτηση μίας πρότασης, πριν από την παρέμβαση της Ελισάβετ.
Ελισάβετ	Κριτικός σκεπτικιστής	Αμφισβητεί παραδοχές, εξετάζει εξαιρέσεις και αναδεικνύει εναλλακτικές οπτικές ή ακραίες περιπτώσεις.	Μία ερώτηση μίας πρότασης, μετά τη Ζωή και την απάντηση του διδάσκοντος.

Πίνακας 3.3: Οι Τρεις Ρόλοι του Συστήματος

Πριν από κάθε σωκρατική απάντηση, το prompt ζητά τρεις αυτοελέγχους: αν η απάντηση αντιφάσκει με όσα είπε ο Κύριος Γιώργος στο ίδιο turn (γύρο διαλόγου), αν περιέχει ασύμβατες δηλώσεις και αν στηρίζεται στη γνωστική βάση. Αν κάποιος έλεγχος αποτύχει, ζητείται αναδιατύπωση πριν από την εμφάνιση. Ο SOCRATIC CONSISTENCY GUARD (έλεγχος συνέπειας του σωκρατικού διαλόγου) αφορά το περιεχόμενο που παράγει το μοντέλο και διαφέρει από τον δείκτη C, που αφορά τις απαντήσεις του μαθητή. Εκτελείται από το ίδιο το μοντέλο, σύμφωνα με τις οδηγίες.

3.6 Κύκλος Λειτουργίας

Η συνεδρία αρχίζει όταν ο χρήστης εισαγάγει username και επιλέξει βιβλίο και κεφάλαιο. Η διεπαφή δημιουργεί τοπικό session_id, μηδενίζει τη μεταβλητή messages και στέλνει το αρχικό μήνυμα «ξεκινάμε». Το prompt ζητά τάξη και αριθμό ερωτήσεων πριν από τη δημιουργία του κουίζ, με προεπιλεγμένο αριθμό πέντε ερωτήσεων.

Ο διακομιστής προσθέτει στο αίτημα τις οδηγίες, τις σημειώσεις και το διαθέσιμο ιστορικό. Οι ερωτήσεις παρουσιάζονται μία κάθε φορά, με τέσσερις επιλογές A/B/C/D, και για καθεμία το μοντέλο εφαρμόζει τον αυτοέλεγχο της Ενότητας 3.2.2.

Μετά την επιλογή ακολουθεί αίτημα αιτιολόγησης, χωρίς να αποκαλυφθεί η ορθότητα της επιλογής μέχρι να δοθεί η εξήγηση.

Το μοντέλο καλείται να αναλύσει εσωτερικά την αιτιολόγηση σε βήματα: βασικές προκειμένες, ενδιάμεσα συμπεράσματα, τελικό συμπέρασμα και τύπο συλλογισμού. Στη συνέχεια αποδίδει τιμές V, L, R και Clarity με βάση 20 ελέγχους ανά αιτιολόγηση. Οι πέντε έλεγχοι του C εφαρμόζονται μόνο στο τέλος, συγκρίνοντας τις απαντήσεις του κουίζ.

Ακολουθούν η εξήγηση, ο ορατός πίνακας βαθμών, η παραγωγή Q_RECORD, ο διάλογος των ρόλων και η ερώτηση κατανόησης. Κάθε φορά που ο μαθητής ζητά επανεξήγηση, δίνεται νέα εξήγηση και σύντομη ερώτηση εφαρμογής, χωρίς νέα βαθμολόγηση. Η ροή συνεχίζεται έως την τελευταία ερώτηση.

Στο τέλος, το μοντέλο υπολογίζει τον C, τους τελικούς Q ανά ερώτηση και τον μέσο Q του κουίζ, παρουσιάζει την τελική αξιολόγηση και παράγει το QUIZ_SUMMARY· οι υπολογισμοί του ελέγχονται με επανυπολογισμό στην Ενότητα 5.6. Το PHP αποθηκεύει μόνιμα τη σύνοψη, ενώ τα αναλυτικά Q_RECORD διατηρούνται στον φυλλομετρητή και παραμένουν διαθέσιμα στο μοντέλο μέσω του ιστορικού συνομιλίας (Ενότητες 3.8 και 3.10).

3.7 Ρουμπρίκα Αξιολόγησης VCLRCI-Q

Το VCLRCI-Q είναι η δομημένη ρουμπρίκα του system prompt για τον μαθητικό συλλογισμό. Προβλέπει 20 ελέγχους ανά αιτιολόγηση, πέντε για κάθε V, L, R και Clarity, και πέντε επιπλέον ελέγχους C στην τελική αποτίμηση. Οι αριθμητικές τιμές V, L, R και Clarity εμφανίζονται στον μαθητή μετά την αιτιολόγηση, ενώ ο C παραμένει σε εκκρεμότητα μέχρι το τέλος. Τα επιμέρους checks (έλεγχοι) και η ανάλυση του συλλογισμού σε βήματα παραμένουν εσωτερικά και, σύμφωνα με το system prompt, δεν εμφανίζονται στον μαθητή.

3.7.1 Οι 20 έλεγχοι ανά αιτιολόγηση

Οι έλεγχοι είναι οδηγίες προς το μοντέλο και όχι ξεχωριστός αλγόριθμος. Το μοντέλο καλείται να εφαρμόσει τους ελέγχους V1–V5, L1–L5, R1–R5 και C11–C15 και να δώσει σε κάθε δείκτη τιμή από 0 έως 1. Η χρήση ενός LLM για κρίσεις βάσει κριτηρίων συνδέεται με τη γενική προσέγγιση LLM-as-Judge (LLM ως αξιολογητής· Zheng et al., 2023), με τη διαφορά ότι εδώ το ίδιο μοντέλο διδάσκει και αξιολογεί.

Οι 20 έλεγχοι οργανώνονται σε τέσσερις ομάδες των πέντε. Κάθε έλεγχος λαμβάνει TRUE (ναι) = 0,2, PARTIAL (μερικώς) = 0,1 ή FALSE (όχι) = 0:

Δείκτης	Check	Ερώτηση που ελέγχει
V	V1	Χρησιμοποιεί σωστό φυσικό νόμο/ορισμό από τις σημειώσεις;
	V2	Οι έννοιες ανήκουν στην ύλη της γνωστικής βάσης;
	V3	Καμία λανθασμένη φυσική δήλωση;
	V4	Σωστό τελικό συμπέρασμα;
	V5	Σωστή αιτιολογική σχέση;
L	L1	Υπάρχει σαφής συλλογισμός βήμα προς βήμα χωρίς άλματα;
	L2	Κάθε βήμα ακολουθεί το προηγούμενο;
	L3	Σωστή χρήση τύπων/σχέσεων από γνωστική βάση;
	L4	Καμία εσωτερική αντίφαση;
	L5	Ο συλλογισμός οδηγεί στο συμπέρασμα;
R	R1	Χρησιμοποιούνται κατάλληλες έννοιες για τη βαθμίδα;
	R2	Αποφεύγει πολύ προχωρημένες έννοιες;

Δείκτης	Check	Ερώτηση που ελέγχει
	R3	Όχι πολύ επιφανειακό για τη βαθμίδα;
	R4	Ακολουθεί σχολική μεθοδολογία;
	R5	Ο συλλογισμός συμφωνεί με τη λογική του σχολικού προγράμματος;
CI	CI1	Δομημένη εξήγηση (αρχή-μέση-τέλος);
	CI2	Κατανοητές προτάσεις;
	CI3	Σωστή φυσική ορολογία;
	CI4	Καμία αμφισημία;
	CI5	Κατανοητό χωρίς επιπλέον πλαίσιο;

Πίνακας 3.4: Οι 20 Έλεγχοι ανά Αιτιολόγηση

Οι CI1–CI5 παρουσιάζονται χωριστά στον Πίνακα 3.5, επειδή εφαρμόζονται μόνο στο τέλος.

3.7.2 Τα C Checks — Consistency στο Τέλος

Ο C υπολογίζεται μετά την τελευταία ερώτηση, βάσει της σύγκρισης των ισχυρισμών σε όλο το διαθέσιμο ιστορικό του κουίζ. Το prompt ζητά από το μοντέλο να χρησιμοποιήσει το εσωτερικό consistency_log και να εφαρμόσει τους πέντε ακόλουθους ελέγχους:

Check	Ερώτηση που ελέγχει
C1	Καμία αντίφαση μεταξύ απαντήσεων σε διαφορετικές ερωτήσεις;
C2	Υπάρχει συνεπής χρήση εννοιών σε όλο το κουίζ;
C3	Δεν άλλαξε αδικαιολόγητα το μοντέλο συλλογισμού;
C4	Ίδιος τύπος → ίδια έννοια παντού;
C5	Δεν αναιρεί προηγούμενα σωστά συμπεράσματα;

Πίνακας 3.5: Τα 5 C Checks για Consistency

Ο C αφορά συνέπεια μεταξύ απαντήσεων και διαφέρει από τον L4, ο οποίος εξετάζει εσωτερική αντίφαση στην ίδια αιτιολόγηση. Ένα σταθερό αλλά επιστημονικά λανθασμένο μοντέλο μπορεί να έχει υψηλό C. Η εναλλαγή σωστών και λανθασμένων επιλογών δεν αρκεί

για να τεκμηριώσει αντίφαση, ενώ μια αιτιολογημένη αναθεώρηση πρέπει να διακρίνεται από αδικαιολόγητη αλλαγή συλλογισμού.

3.7.3 Υπολογισμός Δεικτών Αξιολόγησης

Σύμφωνα με το prompt, κάθε δείκτης ορίζεται ως άθροισμα των πέντε ελέγχων του (V1–V5, L1–L5, R1–R5 και C11–C15) και παίρνει τιμές από 0 έως 1 με βήμα 0,1. Ο δείκτης C ορίζεται με τον ίδιο τρόπο από τους ελέγχους C1–C5, αλλά υπολογίζεται στο τέλος του κουίζ, επειδή συγκρίνει τις απαντήσεις μεταξύ τους. Όπως σημειώνεται στην Ενότητα 3.1.3, οι επιμέρους έλεγχοι δεν καταγράφονται ως χωριστές τιμές και όλες οι κρίσεις παράγονται από το ίδιο μοντέλο· γι' αυτό οι δείκτες δεν μπορούν να θεωρηθούν στατιστικά ανεξάρτητοι.

Το prompt ζητά να μη συμπληρώνεται με εικασίες ένας ελλιπής συλλογισμός, αλλά να χρησιμοποιείται PARTIAL όπου το κριτήριο δεν ικανοποιείται πλήρως.

3.7.4 Ο Συνολικός Δείκτης Q

Μετά την ολοκλήρωση του κουίζ, κάθε αιτιολόγηση αποτιμάται με τον δείκτη Q_n , ο οποίος συνδυάζει τους δείκτες V, L, R και Cl της συγκεκριμένης ερώτησης με τον δείκτη C της συνεδρίας:

$$Q_n = 0,45 \cdot V_n + 0,20 \cdot C + 0,20 \cdot L_n + 0,10 \cdot R_n + 0,05 \cdot Cl_n$$

όπου V_n , L_n , R_n και Cl_n είναι οι δείκτες της ερώτησης n, ενώ ο C υπολογίζεται μία φορά για όλη τη συνεδρία και χρησιμοποιείται με την ίδια τιμή σε όλες τις ερωτήσεις.

Η τελική τιμή Q της συνεδρίας είναι ο αριθμητικός μέσος των τιμών Q_n :

$$Q = (Q_1 + Q_2 + \dots + Q_N) / N$$

όπου N είναι ο αριθμός των ερωτήσεων του κουίζ.

Q	Χαρακτηρισμός
0,90–1,00	Άριστη Απόδοση
0,80–0,89	Πολύ Καλή Απόδοση
0,70–0,79	Καλή Απόδοση
0,60–0,69	Μέτρια Απόδοση
0,00–0,59	Χρειάζεται Βελτίωση

Πίνακας 3.6: Περιγραφικοί Χαρακτηρισμοί του Τελικού Q

Οι χαρακτηρισμοί είναι αυτοί που ορίζει το system prompt για την τελική αξιολόγηση και χρησιμοποιούνται περιγραφικά, όχι ως επικυρωμένο μέτρο μαθησιακής κατανόησης.

3.7.5 Επιλογή Συντελεστών Βαρύτητας

Οι συντελεστές αποτελούν σχεδιαστική επιλογή: ο V λαμβάνει 45%, οι L και C από 20%, ο R 10% και ο Cl 5%. Το μεγαλύτερο βάρος δίνεται στην επιστημονική ορθότητα (V), ενώ οι L και C αφορούν την αλληλουχία και τη συμβατότητα των ισχυρισμών. Ο R, που εξετάζει την καταλληλότητα της αιτιολόγησης για τη βαθμίδα (Ενότητα 2.5), έχει συμπληρωματικό ρόλο. Ο Cl έχει το μικρότερο βάρος, ώστε μια σαφής αλλά επιστημονικά λανθασμένη απάντηση να μην αποκτά υψηλή συνολική τιμή. Οι συντελεστές δεν προέκυψαν από στατιστική βελτιστοποίηση· ο Q αξιοποιείται περιγραφικά, ως συνοπτική ένδειξη της αξιολόγησης.

3.8 Δομημένα Σήματα Καταγραφής (Backend Signals)

Το prompt ζητά δύο τμήματα JSON μέσα σε ειδικούς οριοθέτες (delimiters), δηλαδή σημάνσεις αρχής και τέλους, όπως `%%Q_RECORD_START%%` και `%%Q_RECORD_END%%`. Η JavaScript εξάγει τα Q_RECORD και το PHP εξάγει το QUIZ_SUMMARY, αφαιρώντας τα αντίστοιχα τμήματα από το εμφανιζόμενο κείμενο. Ο κώδικας ελέγχει μόνο ότι το JSON διαβάζεται σωστά· δεν ελέγχει τα πεδία του, τους βαθμούς ή τη συμφωνία τους με όσα βλέπει ο μαθητής.

3.8.1 Δομημένη Καταγραφή Αξιολόγησης (Q_RECORD) Ανά Ερώτηση

Το Q_RECORD παράγεται μετά τον πίνακα αξιολόγησης και πριν από τον σωκρατικό διάλογο. Περιέχει αριθμό και κείμενο ερώτησης, επιλογές, σωστή και επιλεγμένη απάντηση, is_correct (ορθότητα επιλογής), αυτούσια αιτιολόγηση, summary (σύντομη σύνοψη), scores (βαθμούς), score_analysis (λεκτική αιτιολόγηση των βαθμών), q_score (τιμή Q) και re_explain (ένδειξη επανεξήγησης). Οι V, L, R και Clarity έχουν αριθμητικές τιμές, ενώ τα C, q_score και re_explain καταγράφονται ως "pending" (σε εκκρεμότητα), επειδή η εγγραφή παράγεται πριν από το τέλος του κουίζ και πριν από την ερώτηση κατανόησης.

Το summary περιορίζεται στην απάντηση του μαθητή, στη βασική λογική της αιτιολόγησης του και στην ανάλυση σωστού ή λάθους. Το score_analysis δίνει τιμή και σύντομη λεκτική αιτιολόγηση ανά δείκτη, με αναφορά στους σχετικούς ελέγχους.

Η JavaScript αποθηκεύει κάθε αναγνώσιμο Q_RECORD στο localStorage και αντικαθιστά προηγούμενη εγγραφή με ίδιο q_num. Τα Q_RECORD περιλαμβάνονται επίσης στις απαντήσεις του μοντέλου που διατηρούνται στη μεταβλητή messages, οπότε το μοντέλο έχει πρόσβαση σε αυτά μέσω του ιστορικού σε κάθε επόμενο αίτημα.

3.8.2 Σύνοψη Συνεδρίας (QUIZ_SUMMARY)

Το QUIZ_SUMMARY παράγεται μετά την τελική αξιολόγηση και περιλαμβάνει το θέμα (topic), την τάξη (grade), τον αριθμό ερωτήσεων (num_questions), τις σωστές απαντήσεις (correct_answers), τους μέσους όρους των δεικτών (avg_V, avg_L, avg_R, avg_Clarity), τον δείκτη συνέπειας με τους ελέγχους του (C_score, C1–C5), τον μέσο Q (avg_Q), τα δυνατά και τα αδύνατα σημεία (strong_points, weak_points) και τις ερωτήσεις με επανεξήγηση (re_explain_questions).

Το PHP αποθηκεύει μόνιμα το QUIZ_SUMMARY όταν αυτό διαβάζεται σωστά ως JSON. Για τον περιορισμό του όγκου των αποθηκευμένων δεδομένων, η μόνιμη αποθήκευση περιορίζεται στη σύνοψη. Ο κώδικας διατηρεί ωστόσο κενό πεδίο questions, ώστε, όταν χρειάζεται για ερευνητική χρήση, να μπορούν να αποθηκεύονται και οι αναλυτικές εγγραφές του κουίζ.

Οι αποθηκευμένες συνόψεις τροφοδοτούν την επιλογή δυσκολίας σε επόμενα κουίζ (Ενότητα 3.4). Παραδείγματα των δομών Q_RECORD και QUIZ_SUMMARY παρατίθενται στο Παράρτημα Δ.

3.9 Μηχανισμός Επανεξήγησης (Re-explain Flow)

Το σύστημα περιλαμβάνει μηχανισμό επανεξήγησης, ο οποίος ενεργοποιείται όταν ο μαθητής δηλώσει ότι δεν κατανόησε την προηγούμενη εξήγηση ή ζητήσει περαιτέρω διευκρινίσεις. Σκοπός του είναι η υποστήριξη της κατανόησης χωρίς απλή επανάληψη της ίδιας πληροφορίας.

Το prompt ζητά να εντοπιστεί το σημείο δυσκολίας βάσει των ελέγχων που αξιολογήθηκαν ως FALSE ή PARTIAL και να δοθεί στοχευμένη εξήγηση με διαφορετική αναλογία και απλούστερη γλώσσα. Επειδή οι επιμέρους έλεγχοι δεν διατηρούνται ως τιμές (Ενότητα 3.1.3), το μοντέλο βασίζεται στις λεκτικές αιτιολογήσεις του Q_RECORD που βρίσκονται στο ιστορικό. Η διαδικασία θυμίζει τη σταδιακή υποστήριξη της μάθησης (scaffolding).

Ακολουθεί μία σύντομη ερώτηση εφαρμογής σε απλό σενάριο. Η απάντηση του μαθητή εξετάζεται παιδαγωγικά, χωρίς νέα βαθμολόγηση· δεν αλλάζει τους δείκτες που έχουν ήδη αποδοθεί και δεν δημιουργεί νέο Q_RECORD.

Η επανεξήγηση δίνεται μόνο όταν τη ζητήσει ο μαθητής· το σύστημα δεν επαναλαμβάνει την εξήγηση με δική του πρωτοβουλία. Κάθε φορά που ο μαθητής δηλώνει ξανά ότι δεν κατάλαβε, ακολουθεί νέα επανεξήγηση με ακόμη απλούστερη διατύπωση. Όταν ο μαθητής δηλώσει ότι κατάλαβε, η ροή συνεχίζεται στην επόμενη ερώτηση. Η λειτουργία αυτή καθοδηγείται από τις οδηγίες του prompt. Στα σενάρια του Κεφαλαίου 4 ο μηχανισμός δεν ενεργοποιήθηκε (Ενότητα 5.5.3).

3.10 Αποθήκευση Δεδομένων

Η εφαρμογή διαχωρίζει την τοπική διατήρηση των αναλυτικών δεδομένων από τη μόνιμη αποθήκευση των συνοπτικών αποτελεσμάτων στον διακομιστή. Τα Q_RECORD αποθηκεύονται στο localStorage του φυλλομετρητή, το οποίο διατηρεί τα δεδομένα και μετά το κλείσιμο του φυλλομετρητή, μέχρι να αφαιρεθούν.

Στον διακομιστή, το PHP διατηρεί στον κατάλογο logs/physics_quiz/history/ ένα αρχείο ανά username, μετά την εφαρμογή της συνάρτησης safe_name, που αφαιρεί μη επιτρεπτούς χαρακτήρες από το όνομα. Κάθε εγγραφή περιλαμβάνει session_id που δημιουργεί ο διακομιστής, timestamp (χρονική σήμανση), username, summary και το πεδίο questions για ερευνητική χρήση (Ενότητα 3.8.2). Οι τρεις πιο πρόσφατες συνόψεις παρέχονται στο μοντέλο για την επιλογή δυσκολίας (Ενότητα 3.4) και επιτρέπουν περιγραφικές συγκρίσεις προηγούμενων κουίζ. Το πεδίο username δέχεται ελεύθερο κείμενο και δεν απαιτεί ούτε ψευδώνυμο ούτε ταυτοποίηση του χρήστη.

Η JavaScript προβλέπει εκκαθάριση του localStorage στο τέλος του κουίζ μέσω της συνάρτησης lsClear, όταν επιστραφεί το σήμα quiz_finalized. Το PHP όμως επιστρέφει το saved και όχι το quiz_finalized, οπότε οι εγγραφές παραμένουν στον φυλλομετρητή. Σε κάθε νέο κουίζ δημιουργείται νέο session_id με δικό του κλειδί, χωρίς διαγραφή των προηγούμενων. Δεν υπάρχει επίσης μηχανισμός επαναφοράς της συνομιλίας μετά από ανανέωση της σελίδας.

3.11 Τεχνικοί Περιορισμοί του Συστήματος

Οι κανόνες του system prompt δίνουν κοινή βάση αξιολόγησης για όλες τις απαντήσεις· η σταθερότητα των βαθμών σε επαναλαμβανόμενες εκτελέσεις δεν εξετάστηκε στην παρούσα μελέτη. Ο κώδικας δεν περιλαμβάνει βήμα έγκρισης των βαθμών από εκπαιδευτικό. Ο Πίνακας 3.7 συνοψίζει τους τεχνικούς περιορισμούς και τις δυνατότητες βελτίωσης.

Περιορισμός	Περιγραφή	Μετρίαστικός Παράγοντας
Αστάθεια αποκρίσεων του LLM	Η ίδια απάντηση μπορεί να βαθμολογηθεί διαφορετικά σε επανάληψη· στη μελέτη κάθε σενάριο εκτελέστηκε μία φορά.	Η κοινή ρουμπρίκα καθοδηγεί την κρίση· η σταθερότητα μπορεί να εξεταστεί με δοκιμές επαναληψιμότητας.
Απουσία ανεξάρτητου αξιολογητή	Το ίδιο γλωσσικό μοντέλο διδάσκει και αξιολογεί τον μαθητή.	Μελλοντική σύγκριση με εκπαιδευτικούς ή διαφορετικό μοντέλο.
Εξάρτηση από το γλωσσικό μέσο	Η αξιολόγηση βασίζεται μόνο σε γραπτές απαντήσεις.	Μελλοντική ενσωμάτωση εικόνων και διαγραμμάτων.
Έκταση σημειώσεων	Το μοντέλο λαμβάνει έως 20.000 χαρακτήρες σημειώσεων ανά κεφάλαιο.	Έλεγχος ότι το υλικό καλύπτει τη διδακτική ενότητα.
Αριθμητική του μοντέλου	Οι μέσοι όροι και ο Q υπολογίζονται από το μοντέλο· στα σενάρια ελέγχθηκαν με επανυπολογισμό (Ενότητα 5.6)	Υπολογισμός των μέσων όρων και του Q από τον κώδικα.

Πίνακας 3.7: Τεχνικοί Περιορισμοί και Μετρίαστικοί Παράγοντες

3.12 Σύνοψη Κεφαλαίου

Το κεφάλαιο παρουσίασε την αρχιτεκτονική του συστήματος, την παροχή σημειώσεων, το caching, την προσαρμογή της δυσκολίας βάσει συνόψεων, τους τρεις ρόλους, τον μηχανισμό επανεξήγησης και τη ρουμπρίκα VCLRCI-Q. Διευκρινίστηκε ότι οι αυτοέλεγχοι και η βαθμολόγηση εκτελούνται από το ίδιο μοντέλο βάσει οδηγιών, καθώς και ότι τα αναλυτικά Q_RECORD διατηρούνται τοπικά, ενώ στον διακομιστή αποθηκεύεται μόνιμα η σύνοψη του κουίζ. Το επόμενο κεφάλαιο περιγράφει τη μεθοδολογία των συνθετικών σεναρίων και τις πηγές των δεδομένων που αξιοποιούνται στην ανάλυση.

4. Μεθοδολογία Έρευνας

Η παρούσα μελέτη εξετάζει τη λειτουργία του Ph-D A.I. μέσω τριών συνθετικών σεναρίων αλληλεπίδρασης. Η κύρια προσέγγιση είναι ποιοτική και συμπληρώνεται από περιγραφική αριθμητική ανάλυση των παραγόμενων δεικτών. Η χρήση συνθετικών σεναρίων επιτρέπει τον έλεγχο της λειτουργίας του συστήματος με προκαθορισμένα και πλήρως γνωστά πρότυπα απαντήσεων. Οι απαντήσεις διαμορφώθηκαν και εισήχθησαν από τον ερευνητή, χωρίς συμμετοχή πραγματικών μαθητών.

Τα σενάρια εκτελέστηκαν στη διαδικτυακή εφαρμογή και οι διάλογοι παρατίθενται στα Παραρτήματα Α–Γ. Ο σχεδιασμός τους αξιοποιεί ως θεωρητικό υπόβαθρο τη βιβλιογραφία για τις εναλλακτικές αντιλήψεις στη μηχανική (Driver et al., 1985· Viennot, 1979· Galili, 1993). Τα δεδομένα επιτρέπουν την περιγραφή των αποκρίσεων του συστήματος στις συγκεκριμένες εισόδους.

4.1 Ερευνητικά Ερωτήματα

Η μεθοδολογία οργανώθηκε με βάση τα τέσσερα ερευνητικά ερωτήματα της Ενότητας 1.2. Για τις ανάγκες της ανάλυσης, κάθε ερώτημα συνδέθηκε με συγκεκριμένη εστίαση και αντίστοιχες πηγές δεδομένων, όπως παρουσιάζεται στον Πίνακα 4.1.

Ερώτημα	Εστίαση	Πηγή και τρόπος ανάλυσης
EE1: Με ποιον τρόπο το διαλογικό σύστημα αναγνωρίζει και αποτυπώνει προκαθορισμένες παρανοήσεις στις αιτιολογήσεις των συνθετικών σεναρίων;	Αναγνώριση παρανοήσεων	Αιτιολογήσεις και αντίστοιχες εξηγήσεις στους διαλόγους.
EE2: Ποια χαρακτηριστικά εμφανίζει η παραγόμενη ανατροφοδότηση και πώς συνδέονται με τις αρχές της εκπαιδευτικής ανατροφοδότησης των Hattie και Timperley (2007);	Χαρακτηριστικά ανατροφοδότησης	Περιγραφική ταξινόμηση των μηνυμάτων ανατροφοδότησης στα τέσσερα επίπεδα των Hattie και Timperley.

Ερώτημα	Εστίαση	Πηγή και τρόπος ανάλυσης
EE3: Πώς αξιοποιούνται οι δείκτες Validity (V, εγκυρότητα), Consistency (C, συνέπεια), Logic (L, λογική), Regulation (R, εκπαιδευτική καταλληλότητα) και Clarity (Cl, σαφήνεια) για την περιγραφική και συγκριτική ανάλυση των απαντήσεων και της αιτιολόγησης στα τρία συνθετικά σενάρια;	Χρήση δεικτών	Τιμές των δεικτών ανά ερώτηση και τελική αξιολόγηση του συστήματος (Παραρτήματα Α–Γ).
EE4: Πώς οργανώνεται η διαλογική αλληλεπίδραση μέσω των τριών παιδαγωγικών ρόλων και ποιες δυνατότητες αιτιολόγησης και αναστοχασμού δημιουργεί ο σχεδιασμός της;	Οργάνωση διαλόγου	Ακολουθία ρόλων, ερωτήματα και αντίδραση στην αιτιολόγηση.

Πίνακας 4.1: Ερευνητικά Ερωτήματα, Πηγές και Τρόπος Ανάλυσης

4.2 Ερευνητικός Σχεδιασμός — Μελέτη Περίπτωσης

Η έρευνα οργανώνεται ως μελέτη περίπτωσης (case study), σύμφωνα με τον Yin (2018). Η «περίπτωση» που μελετάται είναι ένα σύστημα, το Ph-D A.I., όπως λειτουργεί στη διαδικτυακή εφαρμογή. Μέσα σε αυτό εξετάζονται χωριστά μικρότερα τμήματα: τα τρία σενάρια και οι ερωτήσεις κάθε σεναρίου. Γι' αυτό ο σχεδιασμός χαρακτηρίζεται ως ενιαία μελέτη περίπτωσης με ενσωματωμένες μονάδες ανάλυσης (single-case embedded design).

Η μελέτη δεν αφορά σχολική τάξη ή πειραματική αξιολόγηση διδασκαλίας· εξετάζει πώς απαντά το σύστημα, με ερωτήσεις, εξηγήσεις και βαθμούς, στις εισόδους του ερευνητή. Τα βασικά στοιχεία του σχεδιασμού είναι:

1. Φαινόμενο: η απόκριση του συστήματος σε προκαθορισμένα πρότυπα επιλογών και αιτιολογήσεων.
2. Πλαίσιο: η διαδικτυακή εφαρμογή με το system prompt (τις βασικές οδηγίες προς το μοντέλο), τις σχολικές σημειώσεις και τον κώδικα HTML/JavaScript και PHP.
3. Μονάδες ανάλυσης: τρία συνθετικά προφίλ, με πέντε ερωτήσεις το καθένα, και οι επιμέρους κύκλοι επιλογής, αιτιολόγησης και ανατροφοδότησης.

Ο σχεδιασμός αυτός επιτρέπει τη λεπτομερή περιγραφή συγκεκριμένων αλληλεπιδράσεων. Δεν περιλαμβάνει τυχαιοποίηση, ομάδα ελέγχου ή μέτρηση πριν και μετά τη διδασκαλία· γι' αυτό τα συμπεράσματα είναι περιγραφικά και αφορούν τη λειτουργία του συστήματος, όχι την επίδρασή του στη μάθηση.

4.3 Σχεδιασμός Συνθετικών Σεναρίων

Τα δύο πρώτα σενάρια αναπαριστούν διαφορετικά πρότυπα παρανοήσεων, ενώ το τρίτο περιλαμβάνει επιστημονικά ορθές αιτιολογήσεις. Για τη δημιουργία τους επιλέχθηκαν πρώτα οι εννοιολογικές δυσκολίες, έπειτα γράφτηκαν οι συνθετικές απαντήσεις και τέλος εισήχθησαν στην εφαρμογή.

4.3.1 Επιλογή Παρανοήσεων από Βιβλιογραφία

Το σενάριο Μαθητής_A περιλαμβάνει τη σύγχυση μάζας και βάρους και την αντίληψη ότι η μικρότερη τιμή βάρους στη Σελήνη συνεπάγεται μικρότερη μάζα (M_1 , M_2).

Το σενάριο Μαθητής_B περιλαμβάνει την αντίληψη ότι το βάρος είναι σταθερή ιδιότητα του σώματος, ανεξάρτητη από το βαρυτικό πεδίο (M_3). Στις Ερ.4 και 5 οι αιτιολογήσεις εκφράζουν επίσης ταύτιση μάζας και βάρους και μείωση της μάζας στη Σελήνη.

Το σενάριο Μαθητής_Γ περιλαμβάνει επιστημονικά ορθές επιλογές και αιτιολογήσεις και λειτουργεί ως σημείο αναφοράς για την περιγραφική σύγκριση με τα άλλα δύο. Δεν αποτελεί ομάδα ελέγχου, καθώς ο σχεδιασμός δεν περιλαμβάνει πραγματικούς συμμετέχοντες.

Οι συνθετικές διατυπώσεις είναι κατασκευές του ερευνητή με βάση τη βιβλιογραφία της Ενότητας 2.2 (Driver et al., 1985· Viennot, 1979· Galili, 1993).

4.3.2 Σχεδιασμός Απαντήσεων

Οι απαντήσεις σχεδιάστηκαν από τον ερευνητή σε απλό σχολικό λόγο, ώστε να εκφράζουν τα επιλεγμένα εννοιολογικά πρότυπα. Ο σχεδιασμός περιλάμβανε:

1. Σκόπιμα λανθασμένες επιλογές και αιτιολογήσεις όταν το ερώτημα αφορούσε άμεσα την επιλεγμένη παρανόηση.
2. Ορθές, μερικώς ορθές ή δηλωμένα τυχαίες επιλογές σε άλλα ερωτήματα, σύμφωνα με το συνθετικό προφίλ.

4.3.3 Εκτέλεση στο Πραγματικό Σύστημα

Τα σενάρια εισήχθησαν στην εφαρμογή Ph-D A.I. της πλατφόρμας imathia-edu.gr, με διαφορετικό username (όνομα χρήστη) για κάθε προφίλ: mathitis_A, mathitis_B και mathitis_C. Τα ονόματα αυτά χρησιμοποιήθηκαν μόνο για την καταγραφή, καθώς η εφαρμογή δέχεται οποιοδήποτε username.

Σε όλα τα σενάρια δηλώθηκε Β΄ Γυμνασίου, επιλέχθηκε η ενότητα «Μάζα και Βάρος» (gymnasio_1) και ζητήθηκαν πέντε ερωτήσεις. Οι κοινές αυτές επιλογές παρέχουν κοινό πλαίσιο σύγκρισης, ενώ οι ερωτήσεις παράγονται από το σύστημα σε κάθε εκτέλεση. Οι διάλογοι παρατίθενται στα Παραρτήματα Α–Γ και ο τρόπος συλλογής τους περιγράφεται στην Ενότητα 4.4.

4.4 Εργαλεία Συλλογής Δεδομένων

Η βασική πηγή δεδομένων είναι οι διάλογοι των τριών σεναρίων, όπως αντιγράφηκαν από τη διεπαφή και παρατίθενται στα Παραρτήματα Α–Γ. Οι αριθμητικές τιμές αντλούνται από τους πίνακες βαθμών κάθε ερώτησης και από την τελική αξιολόγηση του συστήματος.

4.4.1 Χειροκίνητη Καταγραφή Ορατού Διαλόγου

Για κάθε σενάριο ο ερευνητής αντέγραψε τη συνομιλία από τη διεπαφή. Το αντίγραφο περιλαμβάνει:

1. τις ερωτήσεις πολλαπλής επιλογής (Α/Β/Γ/Δ) με τις επιλογές τους·
2. την επιλογή και την αιτιολόγηση του μαθητή·
3. την ανατροφοδότηση του Κυρίου Γιώργου, με ένδειξη ορθότητας και επεξήγηση·
4. τον σωκρατικό διάλογο (Ζωή → Κύριος Γιώργος → Ελισάβετ → Κύριος Γιώργος)·
5. τον πίνακα βαθμών κάθε ερώτησης·
6. την τελική αξιολόγηση του συστήματος, με τους μέσους όρους, τον C, τους ελέγχους C1–C5 και τον Q.

Τα αντίγραφα αποτελούν τη βάση της ποιοτικής ανάλυσης. Η παράθεση των διαλόγων στα Παραρτήματα επιτρέπει τον έλεγχο του υλικού από τον αναγνώστη.

4.4.2 Δεδομένα ανά Ερώτηση (Q_RECORD)

Τα Q_RECORD (Ενότητα 3.8.1) δεν εμφανίζονται στη διεπαφή και δεν εξήχθησαν για την παρούσα ανάλυση. Η ανάλυση ανά ερώτηση βασίζεται στις επιλογές, τις αιτιολογήσεις και τις τιμές που εμφανίζονται στους διαλόγους. Αντίστοιχα, το αν δόθηκε επανεξήγηση εξετάζεται στον διάλογο και όχι στο πεδίο re_explain, που καταγράφεται αρχικά ως "pending".

4.4.3 Συνοπτικά Δεδομένα Συνεδρίας (QUIZ_SUMMARY)

Το QUIZ_SUMMARY (Ενότητα 3.8.2) αποθηκεύεται στον διακομιστή, ενώ οι ίδιες τιμές εμφανίζονται και στην τελική αξιολόγηση του διαλόγου. Για την ανάλυση χρησιμοποιούνται οι τιμές της τελικής αξιολόγησης, όπως καταγράφονται στα Παραρτήματα Α–Γ και αναλύονται στο Κεφάλαιο 5.

Όλες οι πηγές, δηλαδή οι διάλογοι, οι πίνακες βαθμών και η τελική αξιολόγηση, προέρχονται από το ίδιο μοντέλο και την ίδια εκτέλεση· το στοιχείο αυτό λαμβάνεται υπόψη κατά την ερμηνεία.

4.5 Διαδικασία Ανάλυσης Δεδομένων

Η ανάλυση οργανώνεται σε τρία επίπεδα: περιεχόμενο των διαλόγων, αριθμητικές τιμές των δεικτών και περιγραφική σύγκριση των σεναρίων.

4.5.1 Πρώτο Επίπεδο — Ποιοτική Ανάλυση Διαλόγων

Οι διάλογοι εξετάζονται με βάση τις αιτιολογήσεις που εισήγαγε ο ερευνητής και τις αποκρίσεις του συστήματος. Η εξέταση εστιάζει:

1. στον τρόπο με τον οποίο το μοντέλο αναγνωρίζει και περιγράφει τις προκαθορισμένες παρανοήσεις των συνθετικών αιτιολογήσεων·
2. στα χαρακτηριστικά της εξήγησης και στο επίπεδο της ανατροφοδότησης, σύμφωνα με τα τέσσερα επίπεδα των Hattie και Timperley (2007) (Ενότητα 2.4.1)·
3. στη σειρά και το περιεχόμενο των παρεμβάσεων της Ζωής, της Ελισάβετ και του Κυρίου Γιώργου και στις ευκαιρίες διερεύνησης που δημιουργούν.

Για την ταξινόμηση της ανατροφοδότησης καταγράφεται το κυρίαρχο επίπεδο κάθε μηνύματος, με αναγνώριση ότι μπορεί να συνυπάρχουν περισσότερα. Ένας έπαινος ή μια

ερώτηση κατανόησης δεν αρκεί από μόνος του για να θεωρηθεί ανατροφοδότηση αυτορρύθμισης.

4.5.2 Δεύτερο Επίπεδο — Ανάλυση Δεικτών

Οι αριθμητικές τιμές εξετάζονται περιγραφικά. Οι V, L, R και Clarity συγκρίνονται ανά ερώτηση, ενώ ο C έχει μία τελική τιμή ανά σενάριο. Η διαδικασία περιλαμβάνει:

1. σύγκριση των τιμών ανά ερώτηση σε σχέση με το περιεχόμενο της αιτιολόγησης·
2. ερμηνεία των τιμών με βάση την αιτιολόγηση, χωρίς ανακατασκευή των επιμέρους ελέγχων·
3. εξέταση των ελέγχων C1–C5, με αναζήτηση συγκεκριμένων συμβατών ή ασύμβατων ισχυρισμών μεταξύ των αιτιολογήσεων·
4. έλεγχο του υπολογισμού του Q ανά ερώτηση και του μέσου Q ανά σενάριο με τον τύπο της Ενότητας 3.7.4 και σύγκριση με την τιμή που έδωσε το σύστημα.

Οι δείκτες C και R ερμηνεύονται σύμφωνα με τους ορισμούς της Ενότητας 2.5.

4.5.3 Τρίτο Επίπεδο — Συγκριτική Ανάλυση

Τα σενάρια συγκρίνονται ως προς τις αριθμητικές τιμές, το περιεχόμενο των αιτιολογήσεων και τη μορφή της ανατροφοδότησης. Οι βασικές παράμετροι συνοψίζονται στον Πίνακα 4.2.

Το σενάριο Μαθητής_Γ προσφέρει σημείο αναφοράς για τις αποκρίσεις του συστήματος σε ορθές αιτιολογήσεις. Οι διαφορές του Q μεταξύ των σεναρίων περιγράφουν πώς βαθμολογήθηκαν τα συγκεκριμένα συνθετικά πρότυπα.

Παράμετρος	Τι συγκρίνεται	Όριο ερμηνείας
V, L, R, Clarity	Τιμές ανά ερώτηση και μέσοι όροι.	Πώς βαθμολογήθηκε κάθε συνθετική απάντηση.
C και C1–C5	Τελική τιμή και έλεγχοι συνέπειας.	Τη συνέπεια μεταξύ των απαντήσεων και την παρουσία πιθανών αντιφάσεων ή αδικαιολόγητων μεταβολών του συλλογισμού.
Q	Σταθμισμένες τιμές και μέσος όρος.	Τη συνολική αποτίμηση και την ορθότητα του υπολογισμού της.

Παράμετρος	Τι συγκρίνεται	Όριο ερμηνείας
Εξηγήσεις	Εξηγήσεις, παραδείγματα και ερωτήματα των ρόλων.	Τα χαρακτηριστικά της ανατροφοδότησης του συστήματος.
Αιτιολογήσεις	Ομοιότητες και διαφορές των συνθετικών απαντήσεων.	Πώς ανταποκρίνεται το σύστημα σε διαφορετικά πρότυπα συλλογισμού.

Πίνακας 4.2: Παράμετροι Συγκριτικής Ανάλυσης

4.6 Μεθοδολογικοί Περιορισμοί

Οι μεθοδολογικοί περιορισμοί που πρέπει να ληφθούν υπόψη κατά την ερμηνεία των αποτελεσμάτων συνοψίζονται στον Πίνακα 4.3.

Περιορισμός	Περιγραφή	Επίπτωση στην Ερμηνεία
Συνθετικά σενάρια	Οι απαντήσεις σχεδιάστηκαν από τον ερευνητή· δεν συμμετείχαν πραγματικοί μαθητές.	Τα αποτελέσματα περιγράφουν τη λειτουργία του συστήματος και δεν γενικεύονται σε μαθητικό πληθυσμό· δεν μετράται μάθηση ή εμπλοκή.
Απουσία σύγκρισης	Δεν υπάρχει ομάδα μαθητών ή σύγκριση με άλλη μέθοδο· το Μαθητής_Γ είναι συνθετικό σημείο αναφοράς.	Δεν εκτιμάται η αποτελεσματικότητα του συστήματος σε σχέση με άλλες προσεγγίσεις.
Περιορισμένο γνωστικό πεδίο	Εξετάζεται μόνο η ενότητα «Μάζα και Βάρος» της Β' Γυμνασίου.	Τα αποτελέσματα αφορούν τη συγκεκριμένη ενότητα και βαθμίδα.
Μικρό σύνολο ερωτήσεων	Τρία σενάρια με πέντε ερωτήσεις το καθένα, διαφορετικές σε κάθε εκτέλεση.	Οι διαφορές μεταξύ σεναρίων μπορεί να σχετίζονται και με το περιεχόμενο ή τη δυσκολία των ερωτήσεων.
Προέλευση δεδομένων	Οι διάλογοι αντιγράφηκαν χειροκίνητα από τη διεπαφή.	Η ανάλυση βασίζεται στους βαθμούς κάθε ερώτησης και στην τελική αξιολόγηση (Παραρτήματα Α–Γ).

Περιορισμός	Περιγραφή	Επίπτωση στην Ερμηνεία
Κρίσεις του ίδιου μοντέλου	Το ίδιο LLM παράγει και αξιολογεί, χωρίς ανεξάρτητη βαθμολόγηση.	Η συμφωνία των εξόδων δεν αποτελεί εξωτερική επικύρωση.
Επανεξήγηση	Ο μηχανισμός δεν ενεργοποιήθηκε, καθώς όπου τέθηκε ερώτηση κατανόησης δόθηκε θετική απάντηση.	Η λειτουργία της Ενότητας 3.9 δεν αξιολογήθηκε.
Μία εκτέλεση ανά σενάριο	Κάθε σενάριο εκτελέστηκε μία φορά, χωρίς ρύθμιση της παραμέτρου temperature.	Η σταθερότητα των βαθμών σε επανάληψη δεν εξετάστηκε.

Πίνακας 4.3: Μεθοδολογικοί Περιορισμοί και Επιπτώσεις

Ο ερευνητής, που είναι και δημιουργός του συστήματος, σχεδίασε τα σενάρια, εισήγαγε τις απαντήσεις και ανέλυσε τις εξόδους, χωρίς ανεξάρτητη αξιολόγηση ή τυφλή κωδικοποίηση. Η συνθήκη αυτή εισάγει πιθανότητα ερευνητικής μεροληψίας. Η αξιοπιστία των κρίσεων μπορεί να εξεταστεί σε μελλοντική διερεύνηση με διαφορετικούς αξιολογητές και συστηματικές επαναλήψεις.

4.7 Ηθικά και Δεοντολογικά Ζητήματα

Η μελέτη χρησιμοποίησε συνθετικά προφίλ και τεχνητά usernames και δεν συνέλεξε απαντήσεις ή προσωπικά δεδομένα πραγματικών μαθητών. Από τον κώδικα της εφαρμογής, στην τρέχουσα μορφή της, προκύπτουν τα εξής:

1. Το πεδίο «Όνομα / Username» δέχεται ελεύθερο κείμενο, οπότε η ανωνυμία εξαρτάται από τον χρήστη.
2. Η τελική σύνοψη αποθηκεύεται αυτόματα, χωρίς προηγούμενη ερώτηση συναίνεσης.
3. Τα μηνύματα της συνομιλίας και το username αποστέλλονται στο API του παρόχου του μοντέλου.
4. Δεν έχει οριστεί πρωτόκολλο για την ενημέρωση των χρηστών, τον χρόνο διατήρησης, την πρόσβαση και τη διαγραφή των δεδομένων.

Πριν από εφαρμογή με πραγματικούς μαθητές χρειάζεται δεοντολογικό πρωτόκολλο που να περιλαμβάνει:

- ενημέρωση για τα δεδομένα που συλλέγονται και για τον εξωτερικό πάροχο·
- διαδικασία συναίνεσης και εγκρίσεων, με έγγραφη συναίνεση γονέων ή κηδεμόνων για ανήλικους μαθητές·
- χρήση ερευνητικών αναγνωριστικών αντί για ονόματα·
- καθορισμό πρόσβασης και χρόνου διατήρησης των δεδομένων·
- δυνατότητα αποχώρησης και διαγραφής.

4.8 Αξιοπιστία και Εγκυρότητα

Η αξιοπιστία της μελέτης στηρίζεται στη διαφάνεια της διαδικασίας: περιγράφονται οι συνθετικές είσοδοι, οι οδηγίες αξιολόγησης του system prompt και ο κώδικας της εφαρμογής, ενώ οι πλήρεις διάλογοι παρατίθενται στα Παραρτήματα Α–Γ, ώστε κάθε εύρημα να μπορεί να ελεγχθεί στην αντίστοιχη απόκριση του συστήματος.

Η εγκυρότητα του σχεδιασμού στηρίζεται στη βιβλιογραφία για τις εναλλακτικές αντιλήψεις στη Φυσική, από την οποία προέκυψαν τα πρότυπα των σεναρίων (Driver et al., 1985· Viennot, 1979· Galili, 1993). Η κοινή ρουμπρίκα του system prompt εφαρμόζει τα ίδια κριτήρια σε όλες τις απαντήσεις.

Οι περιορισμοί που αφορούν τη γενίκευση, τη σταθερότητα των βαθμών και την απουσία ανεξάρτητης αξιολόγησης συνοψίζονται στην Ενότητα 4.6. Οι αριθμητικές τιμές ερμηνεύονται ως έξοδοι του πρωτοτύπου στις εξεταζόμενες συνθήκες· το καθεστώς χρήσης τους διευκρινίζεται στην Ενότητα 4.9.

4.9 Καθεστώς Ερευνητικής Χρήσης των Δεικτών

Οι V, L, R, Clarity και C χρησιμοποιούνται ως περιγραφικοί δείκτες, όπως ορίζονται στο system prompt, και ο Q ως ο σταθμισμένος συνδυασμός τους. Οι τιμές τους επιτρέπουν τη σύγκριση των συγκεκριμένων σεναρίων, χωρίς να αποτελούν τυποποιημένες ψυχομετρικές κλίμακες. Η εξωτερική επικύρωση του πλαισίου αποτελεί το επόμενο ερευνητικό βήμα και εξετάζεται στην Ενότητα 6.6.2.

4.10 Σύνοψη Μεθοδολογίας

Η μελέτη εξέτασε τρία συνθετικά σενάρια στην εφαρμογή, με πέντε ερωτήσεις ανά σενάριο. Οι διάλογοι αντιγράφηκαν χειροκίνητα και αναλύονται μαζί με τους βαθμούς κάθε ερώτησης και την τελική αξιολόγηση του συστήματος. Η ανάλυση εστιάζει στην αναγνώριση των προκαθορισμένων παρανοήσεων, στα χαρακτηριστικά της ανατροφοδότησης, στη χρήση των δεικτών και στην ακολουθία των ρόλων, σύμφωνα με τα τέσσερα ερευνητικά ερωτήματα. Τα αποτελέσματα παρουσιάζονται στο Κεφάλαιο 5.

5. Ανάλυση Δεδομένων και Αποτελεσμάτων

Το κεφάλαιο παρουσιάζει τις αποκρίσεις του συστήματος στα τρία συνθετικά σενάρια: τους βαθμούς κάθε ερώτησης, τον έλεγχο του υπολογισμού του Q και τη σύνδεση των ευρημάτων με τα ερευνητικά ερωτήματα. Η ανάλυση βασίζεται στους διαλόγους των Παραρτημάτων Α–Γ, δηλαδή στους βαθμούς κάθε ερώτησης και στην τελική αξιολόγηση του συστήματος (Ενότητα 4.4). Οι διαφορές μεταξύ των ερωτήσεων εξετάζονται ως λειτουργική συμπεριφορά του συστήματος σε διαφορετικές εισόδους.

5.1 Πλαίσιο Σεναρίων

Η ανάλυση περιλαμβάνει τρία συνθετικά σενάρια με πέντε ερωτήσεις το καθένα: δύο σενάρια με διαφορετικές παρανοήσεις για τη μάζα και το βάρος (Μαθητής_Α και Μαθητής_Β) και ένα σενάριο αναφοράς με επιστημονικά ορθές απαντήσεις (Μαθητής_Γ). Τις επιλογές και τις αιτιολογήσεις τις εισήγαγε ο ερευνητής, και ο σχεδιασμός τους παρουσιάστηκε στην Ενότητα 4.3.

Η ορθότητα της ανατροφοδότησης εξετάζεται ως προς τις σχολικές σημειώσεις της Β' Γυμνασίου που παρέχονται στο σύστημα, με τα ίδια κριτήρια που εφαρμόζει η ρουμπρίκα στις αιτιολογήσεις (δείκτες V και R).

5.2 Σενάριο 1 (Μαθητής_Α)

Το σενάριο Μαθητής_Α περιλαμβάνει δύο συνδεδεμένες παρανοήσεις: (M1) «η μάζα αλλάζει στη Σελήνη» και (M2) «μάζα = βάρος». Η μελέτη των Driver et al. (1985) για τις εναλλακτικές αντιλήψεις στη μηχανική παρέχει το θεωρητικό πλαίσιο του σχεδιασμού.

5.2.1 Ανάλυση ανά Ερώτηση

Ερώτηση 1/5 — «Χαρακτηριστικό της μάζας»

Επιλογή μαθητή: D («Παίρνει διαφορετικές τιμές σε διαφορετικούς πλανήτες»)

Εξήγηση μαθητή: «Γιατί στη σελήνη είμαστε ελαφρύτεροι, οπότε μειώνεται η μάζα μας.»

Δείκτης	Τιμή	Περιγραφή	Ερμηνεία της αιτιολόγησης
V	0.10	Επιστημονική ορθότητα	Η αίσθηση ελαφρότητας αποδίδεται λανθασμένα σε μείωση της μάζας.

Δείκτης	Τιμή	Περιγραφή	Ερμηνεία της αιτιολόγησης
L	0.20	Λογική ακολουθία	Διατυπώνεται αιτιολογική σύνδεση, αλλά η φυσική παραδοχή είναι εσφαλμένη.
R	0.50	Εκπαιδευτική καταλληλότητα	Η τιμή αφορά την εκπαιδευτική καταλληλότητα για τη βαθμίδα.
Clarity	0.70	Σαφήνεια	Η παρανόηση διατυπώνεται κατανοητά.

Πίνακας 5.1: Score Analysis (Ανάλυση Βαθμολογίας) Ερ.1 Σεναρίου 1

Παρατήρηση: Η Clarity=0,70 συνυπάρχει με V=0,10. Η διατύπωση αποδίδει καθαρά τη σύγκριση ανάμεσα στην αίσθηση ελαφρότητας και στη μάζα. Αυτό δείχνει ότι η ρουμπρίκα ξεχωρίζει τη σαφήνεια από την επιστημονική ορθότητα.

Ερώτηση 2/5 — «Όργανο μέτρησης μάζας»

Επιλογή: Α (Δυναμόμετρο)

Εξήγηση: «Γιατί όσο πιο βαρύ είναι ένα σώμα τόσο πιο πολύ δύναμη ασκεί.»

Δείκτης	Τιμή	Περιγραφική ανάλυση
V	0.20	Η αιτιολόγηση συνδέει τη βαρύτητα με τη δύναμη, αλλά δεν διακρίνει το κατάλληλο όργανο μέτρησης μάζας.
L	0.30	Υπάρχει σύνδεση βάρους και δύναμης, χωρίς τεκμηρίωση της επιλογής δυναμόμετρου ως οργάνου μέτρησης μάζας.
R	0.50	Η τιμή αφορά την εκπαιδευτική καταλληλότητα.
Clarity	0.80	Η σύνδεση βάρους και δύναμης διατυπώνεται κατανοητά, παρότι η επιλογή χαρακτηρίζεται λανθασμένη.

Πίνακας 5.2: Score Analysis Ερ.2 Σεναρίου 1

Παρατήρηση: Η ανατροφοδότηση διακρίνει σωστά το δυναμόμετρο, που μετρά το βάρος, από τον ζυγό σύγκρισης, που μετρά τη μάζα. Η αναφορά ότι ο ζυγός λειτουργεί «και στο διάστημα» ισχύει μόνο όπου υπάρχει βαρυτικό πεδίο.

Ερώτηση 3/5 — «Τύπος επιμήκυνσης ελατηρίου»

Επιλογή: Α (τυχαία)

Εξήγηση: «Στην τύχη.»

Δείκτης	Τιμή	Περιγραφική ανάλυση
V	0.00	Δεν παρέχεται φυσική αιτιολόγηση της τυχαίας επιλογής.
L	0.00	Δεν αναπτύσσεται ακολουθία επιχειρημάτων.
R	0.50	Η εκπαιδευτική καταλληλότητα αξιολογείται ανεξάρτητα από τη δυσκολία της ερώτησης.
Clarity	0.50	Η δήλωση άγνοιας είναι κατανοητή, αλλά δεν περιέχει δομημένη εξήγηση.

Πίνακας 5.3: Score Analysis Ερ.3 Σεναρίου 1

Παρατήρηση: Η δήλωση «Στην τύχη» δεν περιέχει φυσικό συλλογισμό. Η ανατροφοδότηση του συστήματος εξηγεί τη σχέση μεταξύ βάρους και επιμήκυνσης του ελατηρίου.

Ερώτηση 4/5 — «Υπολογισμός βάρους»

Επιλογή: A (4N)

Εξήγηση: «Γιατί μάζα και Βάρος είναι το ίδιο.»

Δείκτης	Τιμή	Περιγραφική ανάλυση
V	0.00	Η δήλωση «μάζα και βάρος είναι το ίδιο» είναι επιστημονικά εσφαλμένη.
L	0.10	Η επιλογή υποστηρίζεται από την εσφαλμένη ταύτιση δύο διαφορετικών μεγεθών.
R	0.50	Η τιμή αφορά καταλληλότητα για τη βαθμίδα.
Clarity	0.60	Η απάντηση παρουσιάζει μερική σαφήνεια στη διατύπωση, ωστόσο χρησιμοποιείται λανθασμένη επιστημονική ορολογία.

Πίνακας 5.4: Score Analysis Ερ.4 Σεναρίου 1

Παρατήρηση: Η αιτιολόγηση εκφράζει ρητά την παρανόηση M2, καθώς ταυτίζει μάζα και βάρος, και η τιμή $V=0,00$ συνδέεται με αυτή την εσφαλμένη δήλωση. Το εκφώνημα δίνει $m = 4 \text{ kg}$ και $g = 10 \text{ m/s}^2$, άρα το βάρος είναι 40 N (επιλογή C)· η επιλογή A (4 N) προκύπτει ακριβώς από την ταύτιση μάζας και βάρους.

Ερώτηση 5/5 — «Ίδια μάζα → ίδιο βάρος»

Επιλογή: B (σωστή)

Εξήγηση: «Αφού έχουν ίδια μάζα, θα έχουν και ίδιο βάρος.»

Δείκτης	Τιμή	Περιγραφική ανάλυση
V	1.00	Η αιτιολόγηση εφαρμόζει σωστά τη σχέση ίδιας μάζας και ίδιου βάρους για σώματα στον ίδιο τόπο, όπως ορίζει το εκφώνημα («βρίσκονται στη Γη»).
L	0.90	Σύντομος αλλά επαρκής συλλογισμός για το συγκεκριμένο ερώτημα.
R	0.80	Η απάντηση είναι συμβατή με το γνωστικό επίπεδο της βαθμίδας.
Clarity	0.90	Η αιτιολόγηση είναι σύντομη και κατανοητή.

Πίνακας 5.5: Score Analysis Ερ.5 Σεναρίου 1

Παρατήρηση: Το εκφώνημα ορίζει ότι τα δύο σώματα βρίσκονται στη Γη, επομένως η σύντομη αιτιολόγηση επαρκεί για το συγκεκριμένο ερώτημα. Στην ίδια απόκριση ο Κύριος Γιώργος χρησιμοποιεί ενδεικτική τιμή μάζας («0.2kg») που δεν δίνεται στο εκφώνημα, ενώ η απάντηση στην Ελισάβετ για το g αφορά σημεία κοντά στην επιφάνεια της Γης.

5.2.2 Τιμές Δεικτών ανά Ερώτηση Σεναρίου 1

Ερ.	V	L	R	Clarity	Σωστή;	Παρατήρηση
Ερ.1	0.1	0.2	0.5	0.7	Όχι	Παρανόηση M1 — σαφής αλλά λάθος
Ερ.2	0.2	0.3	0.5	0.8	Όχι	Παρανόηση M2 — συγχέει βάρος/μάζα
Ερ.3	0.0	0.0	0.5	0.5	Όχι	Άγνοια — τυχαία επιλογή
Ερ.4	0.0	0.1	0.5	0.6	Όχι	Παρανόηση M2 στην πιο καθαρή μορφή
Ερ.5	1.0	0.9	0.8	0.9	Ναι	Ορθή απάντηση σε συγκεκριμένο πλαίσιο
Μέσος	0.26	0.30	0.56	0.70	1/5	

Πίνακας 5.6: Τιμές Δεικτών Σεναρίου 1

5.2.3 C Score — Consistency Analysis (Ανάλυση Συνέπειας)

C Check	Τιμή	Έλεγχος ερμηνείας της τιμής
C1	0.1 PARTIAL (μερικώς)	Οι δηλώσεις «μάζα και βάρος είναι το ίδιο» (Ερ.4) και «ίδια μάζα, ίδιο βάρος» (Ερ.5) είναι συμβατές μεταξύ τους· η κρίση PARTIAL εκφράζει τη συνολική εκτίμηση του μοντέλου.
C2	0.1 PARTIAL	Η ιδέα ότι η μάζα αλλάζει (Ερ.1) συμφωνεί με τις υπόλοιπες απαντήσεις του Παραρτήματος Α· η κρίση PARTIAL εκφράζει τη συνολική εκτίμηση του μοντέλου.
C3	0.0 FALSE (όχι)	Στο Παράρτημα Α ο μαθητής κρατά τον ίδιο τρόπο σκέψης (Ερ.4 και Ερ.5)· η κρίση FALSE εκφράζει τη συνολική εκτίμηση του μοντέλου.
C4	0.0 FALSE	Στο Παράρτημα Α κάθε σχέση χρησιμοποιείται με το ίδιο νόημα· η κρίση FALSE εκφράζει τη συνολική εκτίμηση του μοντέλου.
C5	0.2 TRUE (ναι)	Τα σωστά συμπεράσματα διατηρούνται μέχρι το τέλος του κουίζ.

Πίνακας 5.7: C Checks Σεναρίου 1

$$C_score = 0.1+0.1+0.0+0.0+0.2 = 0.40$$

Ερμηνεία: Το άθροισμα των C1–C5 δίνει C = 0,40, που χρησιμοποιείται στον υπολογισμό του Q. Οι απαντήσεις του Σεναρίου 1 είναι συμβατές μεταξύ τους, για παράδειγμα η ταύτιση μάζας και βάρους (Ερ.4) συμφωνεί με το «ίδια μάζα, ίδιο βάρος» (Ερ.5). Η τιμή 0,40 εκφράζει επομένως τη συνολική κρίση του μοντέλου για το κουίζ.

5.2.4 Τελικό Q Score Σεναρίου 1

Τύπος: $Q_n = 0,45 \cdot V_n + 0,20 \cdot C + 0,20 \cdot L_n + 0,10 \cdot R_n + 0,05 \cdot Cl_n$ (Ενότητα 3.7.4), με την ίδια τελική τιμή C = 0,40 σε όλες τις ερωτήσεις. Τα γινόμενα του Πίνακα 5.8 υπολογίζονται από τις τιμές του Πίνακα 5.6.

Ερώτηση	0.45×V	0.20×C	0.20×L	0.10×R	0.05×Cl	Q _n
Q1	0.045	0.080	0.040	0.050	0.035	0.250
Q2	0.090	0.080	0.060	0.050	0.040	0.320

Ερώτηση	0.45×V	0.20×C	0.20×L	0.10×R	0.05×CI	Q _n
Q3	0.000	0.080	0.000	0.050	0.025	0.155
Q4	0.000	0.080	0.020	0.050	0.030	0.180
Q5	0.450	0.080	0.180	0.080	0.045	0.835
Μέσος						0.35

Πίνακας 5.8: Υπολογισμός Q Score Σεναρίου 1

Μέσο Q = 0,35, όσο έδωσε και το σύστημα.

Ερμηνεία: Το Q είναι χαμηλό κυρίως επειδή οι βαθμοί V και L είναι χαμηλοί και το C είναι 0,40. Μόνο η Ερ.5 απαντήθηκε σωστά και έχει Q₅ = 0,835.

5.2.5 Αξιολόγηση Ανατροφοδότησης

Ο Πίνακας 5.9 ταξινομεί τα μηνύματα ανατροφοδότησης του Σεναρίου 1 στα επίπεδα των Hattie και Timperley (2007), σύμφωνα με τη διαδικασία της Ενότητας 4.5.1. Οι περιγραφές συνοψίζουν τα χαρακτηριστικά των μηνυμάτων και δεν αποτελούν αυτούσια αποσπάσματα.

Ερ.	Περιγραφική ταξινόμηση	Σύνοψη χαρακτηριστικών
Ερ.1	Έργο και διαδικασία	Διάκριση μάζας και βάρους και εξήγηση της εσφαλμένης σύνδεσής τους.
Ερ.2	Έργο και διαδικασία	Εξήγηση της λειτουργίας δυναμόμετρου και της εξάρτησης της ένδειξης από το βαρυτικό πεδίο.
Ερ.3	Έργο και διαδικασία	Εξήγηση της σχέσης μεταξύ βάρους και επιμήκυνσης του ελατηρίου.
Ερ.4	Έργο και διαδικασία	Αποσαφήνιση εννοιών, μονάδων και σχέσης μάζας και βάρους.
Ερ.5	Έργο και διαδικασία	Επιβεβαίωση επιλογής και εξήγηση της σχέσης.

Πίνακας 5.9: Τύπος Ανατροφοδότησης ανά Ερώτηση — Σενάριο 1

5.3 Σενάριο 2 (Μαθητής_B)

Το σενάριο Μαθητής_B εκκινεί από την παρανόηση M3: «το βάρος είναι εγγενής ιδιότητα του σώματος και δεν μεταβάλλεται με τον τόπο». Στις μεταγενέστερες αιτιολογήσεις εμφανίζονται επίσης ταύτιση μάζας και βάρους και ο ισχυρισμός ότι και τα δύο μειώνονται

στη Σελήνη. Οι Viennot (1979) και Galili (1993) παρέχουν το ευρύτερο θεωρητικό πλαίσιο για τέτοιες εναλλακτικές αντιλήψεις.

5.3.1 Ανάλυση ανά Ερώτηση

Ερώτηση 1/5 — «Τι είναι σταθερό παντού;»

Επιλογή: A (Βάρος)

Εξήγηση: «Το βάρος δεν αλλάζει, είναι το ίδιο παντού.»

Δείκτης	Τιμή	Περιγραφική ανάλυση
V	0.00	Η αιτιολόγηση δηλώνει ότι το βάρος παραμένει σταθερό παντού.
L	0.10	Ο ισχυρισμός είναι σαφής, αλλά δεν αναπτύσσεται φυσική τεκμηρίωση.
R	0.80	Η τιμή αφορά την καταλληλότητα εννοιών και βάθους για τη βαθμίδα.
Clarity	0.90	Η εσφαλμένη δήλωση εκφράζεται ξεκάθαρα και χωρίς γλωσσική αμφισημία.

Πίνακας 5.10: Score Analysis Ερ.1 Σεναρίου 2

Παρατήρηση: Η Clarity=0,90 είναι η υψηλότερη τιμή για λανθασμένη επιλογή στο σύνολο των σεναρίων. Η αιτιολόγηση εκφράζει σαφώς την παρανόηση M3.

Ερώτηση 2/5 — «Ηλεκτρονικός ζυγός στη Σελήνη»

Επιλογή: A (2kg — ίδιο)

Εξήγηση: «Η μάζα είναι 2kg γιατί το βάρος δεν αλλάζει, είναι χαρακτηριστικό του σώματος.»

Δείκτης	Τιμή	Περιγραφική ανάλυση
V	0.20	Η αιτιολόγηση αναφέρει σωστά ότι η μάζα είναι 2 kg, αλλά δεν διακρίνει την ένδειξη του ηλεκτρονικού ζυγού από τη μάζα και στηρίζεται στον εσφαλμένο ισχυρισμό ότι το βάρος δεν αλλάζει.
L	0.20	Η αιτιολόγηση της τιμής μάζας βασίζεται στον εσφαλμένο ισχυρισμό ότι το βάρος δεν αλλάζει.
R	0.80	Η τιμή περιγράφει την εκπαιδευτική καταλληλότητα.

Δείκτης	Τιμή	Περιγραφική ανάλυση
Clarity	0.80	Η πρόταση είναι κατανοητή, αλλά συγχέει τη μάζα με την εξάρτηση του βάρους από τον τόπο.

Πίνακας 5.11: Score Analysis Ερ.2 Σεναρίου 2

Ερώτηση 3/5 — «Σχέση μάζας-επιμήκυνσης»

Επιλογή: C (ευθέως ανάλογη — σωστή)

Εξήγηση: «Όσο βαρύτερο το σώμα τόσο τεντώνει το ελατήριο, άρα το βάρος και η επιμήκυνση είναι ίδια πράγμα.»

Δείκτης	Τιμή	Περιγραφική ανάλυση
V	1.00	Η επιλογή C είναι σωστή και η αιτιολόγηση συνδέει σωστά μεγαλύτερο βάρος με μεγαλύτερη επιμήκυνση.
L	0.90	Η εξήγηση συνδέει μεγαλύτερο βάρος με μεγαλύτερη επιμήκυνση. Η αναλογική σχέση προϋποθέτει ίδιο ελατήριο και ίδιο βαρυτικό πεδίο.
R	1.00	Η απάντηση αξιολογείται ως κατάλληλη για τη βαθμίδα.
Clarity	0.90	Η διατύπωση είναι κατανοητή.

Πίνακας 5.12: Score Analysis Ερ.3 Σεναρίου 2

Παρατήρηση: Η αιτιολόγηση συνδέει σωστά μεγαλύτερο βάρος με μεγαλύτερη επιμήκυνση, και η ανατροφοδότηση εστιάζει στη σχέση $x = mg/k$, που αποδίδει με ακρίβεια την αναλογία μάζας και επιμήκυνσης.

Ερώτηση 4/5 — «Δύο σώματα 5kg — βάρος»

Επιλογή: A (5N)

Εξήγηση: «Έχουν ίδιο βάρος 5kg γιατί η μάζα και το βάρος είναι το ίδιο.»

Δείκτης	Τιμή	Περιγραφική ανάλυση
V	0.00	Η αιτιολόγηση ταυτίζει μάζα και βάρος και χρησιμοποιεί kg για το βάρος.
L	0.00	Δεν παρέχεται τεκμηρίωση μετατροπής μάζας σε βάρος.
R	0.80	Η τιμή αφορά την εκπαιδευτική καταλληλότητα.

Δείκτης	Τιμή	Περιγραφική ανάλυση
Clarity	0.80	Η ταύτιση των δύο μεγεθών είναι κατανοητή ως δήλωση, παρότι επιστημονικά εσφαλμένη.

Πίνακας 5.13: Score Analysis Ερ.4 Σεναρίου 2

Ερώτηση 5/5 — «Ζυγός vs Δυναμόμετρο στη Σελήνη»

Επιλογή: D (και τα δύο μικρότερες τιμές)

Εξήγηση: «Στη Σελήνη όλα είναι πιο ελαφριά — βάρος και μάζα μειώνονται.»

Δείκτης	Τιμή	Περιγραφική ανάλυση
V	0.00	Η μείωση βάρους στη Σελήνη αποδίδεται λανθασμένα και στη μάζα.
L	0.10	Η αίσθηση ελαφρότητας χρησιμοποιείται ως αιτία μεταβολής και των δύο μεγεθών.
R	0.70	Η τιμή αφορά την καταλληλότητα για τη βαθμίδα.
Clarity	0.80	Η εσφαλμένη σύνδεση διατυπώνεται σαφώς.

Πίνακας 5.14: Score Analysis Ερ.5 Σεναρίου 2

Παρατήρηση: Η ανατροφοδότηση διακρίνει σωστά τον ζυγό σύγκρισης από το δυναμόμετρο και εξηγεί ότι στον ζυγό το g απαλείφεται. Η διατύπωση ότι ο ζυγός είναι «αξιόπιστος παντού στο σύμπαν», σε απάντηση για ταξίδια στο διάστημα, ισχύει μόνο όπου υπάρχει βαρυτικό πεδίο.

5.3.2 Τιμές Δεικτών ανά Ερώτηση Σεναρίου 2

Ερ.	V	L	R	Clarity	Σωστή;	Παρατήρηση
Ερ.1	0.0	0.1	0.8	0.9	Όχι	Εμφανίζεται σαφώς η παρανόηση ότι το βάρος αποτελεί σταθερό χαρακτηριστικό του σώματος.
Ερ.2	0.2	0.2	0.8	0.8	Όχι	Η ίδια παρανόηση επεκτείνεται στην ερμηνεία των μετρήσεων του ζυγού.
Ερ.3	1.0	0.9	1.0	0.9	Ναι	Ορθή επιλογή στο συγκεκριμένο ερώτημα· δεν τεκμηριώνει μεταβολή κατανόησης.

Ερ.	V	L	R	Clarity	Σωστή;	Παρατήρηση
Ερ.4	0.0	0.0	0.8	0.8	Όχι	Παρατηρείται επιπλέον σύγχυση μεταξύ μάζας και βάρους.
Ερ.5	0.0	0.1	0.7	0.8	Όχι	Επανεμφανίζονται πολλαπλές εννοιολογικές δυσκολίες σχετικές με τη διάκριση μάζας και βάρους.
Μέσος	0.24	0.26	0.82	0.84	1/5	

Πίνακας 5.15: Τιμές Δεικτών Σεναρίου 2

5.3.3 C Score — Consistency Analysis

C Check	Τιμή	Έλεγχος ερμηνείας της τιμής
C1	0.1 PARTIAL	Ασυμβατότητα εμφανίζεται μεταξύ «βάρος σταθερό παντού» (Ερ.1) και «βάρος μειώνεται στη Σελήνη» (Ερ.5), ενώ η σωστή Ερ.3 είναι συμβατή με τις υπόλοιπες απαντήσεις.
C2	0.1 PARTIAL	Το βάρος χρησιμοποιείται ως σταθερή ιδιότητα (Ερ.1, Ερ.2) και ως μεταβαλλόμενο στη Σελήνη (Ερ.5), γι' αυτό η κρίση είναι PARTIAL.
C3	0.0 FALSE	Στην Ερ.5 ο μαθητής αλλάζει άποψη για το βάρος χωρίς να εξηγεί γιατί, οπότε η κρίση FALSE είναι αιτιολογημένη.
C4	0.0 FALSE	Η αναλογική σχέση της Ερ.3 χρησιμοποιείται σωστά για τις συνθήκες της ερώτησης, η κρίση FALSE εκφράζει τη συνολική εκτίμηση του μοντέλου.
C5	0.2 TRUE	Το σωστό συμπέρασμα της Ερ.3 διατηρείται μέχρι το τέλος του κουίζ.

Πίνακας 5.16: C Checks Σεναρίου 2

$$C_score = 0.1+0.1+0.0+0.0+0.2 = 0.40$$

Ερμηνεία: Το C = 0,40 είναι ίδιο με το Σενάριο 1, αλλά εδώ στηρίζεται σε συγκεκριμένη αντίφαση. Στην Ερ.1 ο μαθητής λέει ότι «το βάρος δεν αλλάζει», ενώ στην Ερ.5 ότι «βάρος και μάζα μειώνονται» στη Σελήνη, χωρίς να εξηγεί την αλλαγή.

5.3.4 Τελικό Q Score Σεναρίου 2

Ερ.	Q _n (υπολογισμός)	Τιμή
Q1	$0.45 \times 0.0 + 0.20 \times 0.40 + 0.20 \times 0.1 + 0.10 \times 0.8 + 0.05 \times 0.9$	0.225
Q2	$0.45 \times 0.2 + 0.20 \times 0.40 + 0.20 \times 0.2 + 0.10 \times 0.8 + 0.05 \times 0.8$	0.330
Q3	$0.45 \times 1.0 + 0.20 \times 0.40 + 0.20 \times 0.9 + 0.10 \times 1.0 + 0.05 \times 0.9$	0.855
Q4	$0.45 \times 0.0 + 0.20 \times 0.40 + 0.20 \times 0.0 + 0.10 \times 0.8 + 0.05 \times 0.8$	0.200
Q5	$0.45 \times 0.0 + 0.20 \times 0.40 + 0.20 \times 0.1 + 0.10 \times 0.7 + 0.05 \times 0.8$	0.210
Μέσος Q		0.36

Πίνακας 5.17: Υπολογισμός Q Score Σεναρίου 2

Μέσο Q = 0,36, όσο έδωσε και το σύστημα.

Ερμηνεία: Το Q είναι χαμηλό κυρίως επειδή οι βαθμοί V και L είναι χαμηλοί και το C είναι 0,40. Μόνο η Ερ.3 απαντήθηκε σωστά και έχει Q₃ = 0,855. Στις Ερ.4 και 5 ο μαθητής επαναλαμβάνει τις λανθασμένες ιδέες του, όπως είχε σχεδιαστεί το σενάριο, και γι' αυτό τα Q₄ και Q₅ είναι χαμηλά.

5.4 Σενάριο 3 (Μαθητής_Γ)

Το σενάριο Μαθητής_Γ περιλαμβάνει επιστημονικά ορθές συνθετικές απαντήσεις και λειτουργεί ως σημείο αναφοράς. Στη συγκεκριμένη εκτέλεση οι ορθές αιτιολογήσεις συνοδεύτηκαν από υψηλές τιμές δεικτών.

5.4.1 Score Analysis Αναφοράς

Ερ.	V	L	R	Clarity	Αιτιολόγηση (συνοπτική)
Ερ.1	1.00	1.00	1.00	1.00	Ο μαθητής διατυπώνει ορθό ορισμό της μάζας και αναγνωρίζει ότι αποτελεί σταθερό χαρακτηριστικό του σώματος.
Ερ.2	1.00	1.00	1.00	1.00	Η απάντηση συνδέει ορθά την επιμήκυνση του ελατηρίου με το βάρος του σώματος και εφαρμόζει σωστά τις σχετικές φυσικές έννοιες.

Ερ.	V	L	R	Clarity	Αιτιολόγηση (συνοπτική)
Ερ.3	1.00	1.00	1.00	1.00	Αναγνωρίζεται ορθά η αναλογική σχέση μεταξύ βάρους και μάζας όταν η βαρυτική επιτάχυνση παραμένει σταθερή.
Ερ.4	1.00	0.90	1.00	0.90	Η απάντηση είναι επιστημονικά ορθή, ωστόσο η τεκμηρίωση του τελικού αποτελέσματος δεν αναπτύσσεται πλήρως.
Ερ.5	1.00	1.00	1.00	1.00	Διακρίνονται με σαφήνεια οι έννοιες της μάζας και του βάρους καθώς και τα αντίστοιχα όργανα μέτρησής τους.
Μέσος	1.00	0.98	1.00	0.98	Υψηλές τιμές στο σενάριο αναφοράς.

Πίνακας 5.18: Score Analysis Σεναρίου 3

Παρατήρηση για την ανατροφοδότηση της Ερ.2: Ο Κύριος Γιώργος υιοθετεί τη λέξη «βαρύτερο» από την ερώτηση της Ζωής και τη διευκρινίζει αμέσως ως «πιο σκληρό», συνδέοντάς τη σωστά με μεγαλύτερο k και μικρότερη επιμήκυνση για ίδιο σώμα και ίδιο g .

5.4.2 C Score Αναφοράς

C Check	Τιμή	Έλεγχος ερμηνείας της τιμής
C1	0.2 TRUE	Στο Παράρτημα Γ οι αιτιολογήσεις δεν περιέχουν ασύμβατους ισχυρισμούς.
C2	0.2 TRUE	Οι έννοιες χρησιμοποιούνται με τον ίδιο τρόπο σε όλες τις απαντήσεις.
C3	0.2 TRUE	Δεν εμφανίζεται αλλαγή μοντέλου συλλογισμού.
C4	0.2 TRUE	Οι αιτιολογήσεις χρησιμοποιούν με τον ίδιο τρόπο τη σχέση βάρους, μάζας και βαρυτικού πεδίου (Ερ.2–5).
C5	0.2 TRUE	Δεν εμφανίζεται αναίρεση προηγούμενων σωστών συμπερασμάτων.

Πίνακας 5.19: C Checks Σεναρίου 3

$C_score = 1,00$. Με τους μέσους όρους του Πίνακα 5.18, ο μέσος Q υπολογίζεται ως $0,45 \times 1,00 + 0,20 \times 1,00 + 0,20 \times 0,98 + 0,10 \times 1,00 + 0,05 \times 0,98 = 0,995 \approx 1,00$. Η στρογγυλοποιημένη τιμή δεν σημαίνει ότι όλα τα επιμέρους σκορ είναι ίσα με 1,00.

5.5 Συγκριτική Ανάλυση Σεναρίων

5.5.1 Σύγκριση Δεικτών

Δείκτης	Σεν.1	Σεν.2	Σεν.3	Παρατήρηση
V	0.26	0.24	1.00	Χαμηλοί μέσοι βαθμοί στα δύο σενάρια παρανοήσεων και υψηλός στο σενάριο αναφοράς.
L	0.30	0.26	0.98	Χαμηλές τιμές στα σενάρια παρανοήσεων και υψηλή στο σενάριο αναφοράς.
R	0.56	0.82	1.00	Υψηλότερη τιμή στο Σεν.2 από το Σεν.1, παρότι και τα δύο περιλαμβάνουν παρανοήσεις.
Clarity	0.70	0.84	0.98	Η κατανοητή διατύπωση μπορεί να συνυπάρχει με επιστημονικό σφάλμα.
C	0.40	0.40	1.00	Συνέπεια ισχυρισμών, χωριστά από την ορθότητά τους.
Q	0.35	0.36	1.00	Σταθμισμένος συνδυασμός και διαφοροποίηση των σεναρίων παρανοήσεων από το σενάριο αναφοράς.

Πίνακας 5.20: Συγκριτική Σύνοψη Δεικτών

5.5.2 Ευρήματα ανά Ερευνητικό Ερώτημα

EE1 Αναγνώριση προκαθορισμένων παρανοήσεων

Στα δύο πρώτα σενάρια οι εξηγήσεις του συστήματος εστιάζουν στη διάκριση μάζας και βάρους και στην εξάρτηση του βάρους από το βαρυτικό πεδίο. Οι μέσοι $V = 0,26$ και $0,24$ συνυπάρχουν με $Clarity = 0,70$ και $0,84$, αντίστοιχα, γεγονός που δείχνει ότι το σύστημα αποδίδει διαφορετικές κρίσεις για την ορθότητα και τη σαφήνεια της ίδιας αιτιολόγησης.

EE2 Χαρακτηριστικά της παραγόμενης ανατροφοδότησης

Η ανατροφοδότηση περιλαμβάνει αποσαφήνιση φυσικών εννοιών, παραδείγματα και επεξήγηση σχέσεων. Στο Σενάριο 1 τα μηνύματα ανατροφοδότησης εστιάζουν στην ίδια την

απάντηση και στον τρόπο σκέψης του μαθητή, όπως περιγράφουν οι Hattie και Timperley (2007) (Πίνακας 5.9).

ΕΕ3 Περιγραφική και συγκριτική χρήση των δεικτών

Ο V αφορά την επιστημονική ορθότητα, ο L τη λογική ακολουθία, ο R την εκπαιδευτική καταλληλότητα για τη βαθμίδα και η Clarity τη σαφήνεια της διατύπωσης· ο C αφορά τη συνέπεια μεταξύ αιτιολογήσεων και ο Q τον σταθμισμένο συνδυασμό τους. Καθώς κάθε δείκτης αποτυπώνει διαφορετική διάσταση της αιτιολόγησης, οι δείκτες ερμηνεύονται συνδυαστικά.

Οι μέσες τιμές του Πίνακα 5.20 επιτρέπουν περιγραφική σύγκριση των τριών εκτελέσεων. Τα δύο σενάρια παρανοήσεων εμφανίζουν σχεδόν ίδιο Q (0,35 και 0,36), διακριτό από το σενάριο αναφοράς (1,00).

ΕΕ4 Οργάνωση της αλληλεπίδρασης και δυνατότητες αναστοχασμού

Η ακολουθία των τριών ρόλων οργανώνει την εξήγηση και τα ερωτήματα επέκτασης ή κριτικού ελέγχου και τηρήθηκε σε όλες τις ερωτήσεις. Ο Κύριος Γιώργος παρέχει διδακτική εξήγηση, η Ζωή σωκρατική διερεύνηση και η Ελισάβετ κριτικό έλεγχο. Η δομή αυτή δημιουργεί δυνατότητες αιτιολόγησης και αναστοχασμού.

5.5.3 Κύρια Ευρήματα

Ο Πίνακας 5.21 συγκεντρώνει τις κύριες παρατηρήσεις για τη λειτουργία του συστήματος στα τρία σενάρια, ανά ερευνητικό ερώτημα, με παραπομπή στους πίνακες του κεφαλαίου και στα Παραρτήματα.

ΕΕ	Εύρημα	Τεκμηρίωση
ΕΕ1, ΕΕ3	Στις ερωτήσεις με σωστή επιλογή το μοντέλο απέδωσε $V = 1,00$, ενώ στις λανθασμένες επιλογές με παρανόηση απέδωσε V από 0,00 έως 0,20.	Πίν. 5.6, 5.15, 5.18
ΕΕ1, ΕΕ3	Η τιμή $C = 0,40$ και η κατανομή C1–C5 ($0,1 \cdot 0,1 \cdot 0 \cdot 0 \cdot 0,2$) είναι ίδιες στα δύο σενάρια παρανοήσεων. Στο Σενάριο 2 υπάρχει συγκεκριμένη αντίφαση (Ερ.1–Ερ.5), ενώ στο Σενάριο 1 η τιμή αντανάκλα τη συνολική κρίση του μοντέλου.	Πίν. 5.7, 5.16

ΕΕ	Εύρημα	Τεκμηρίωση
EE2	Στην ταξινόμηση του Σεναρίου 1 κυριάρχησαν χαρακτηριστικά επιπέδου έργου και διαδικασίας, με αποσαφήνιση εννοιών, σχέσεις ($W = mg$, $x = mg/k$) και παραδείγματα.	Πίν. 5.9
EE2	Η ανατροφοδότηση ακολουθεί τη σχολική διάκριση μάζας και βάρους και τη διάκριση ζυγού σύγκρισης και δυναμόμετρου.	Παραρτήματα Α–Γ
EE3	Το ίδιο ποσοστό ορθών επιλογών (1/5) στα Σενάρια 1 και 2 συνοδεύεται από διαφορετικά προφίλ R (0,56 και 0,82) και Clarity (0,70 και 0,84).	Πίν. 5.20
EE4	Η ακολουθία των ρόλων τηρήθηκε σε όλες τις ερωτήσεις. Η Ζωή συνέδεσε σε μία περίπτωση τη συζήτηση με την καθημερινή γλώσσα («στην καθημερινή γλώσσα λέμε "ζυγίζω 50 kg" — αυτό δεν είναι λάθος;»). Στο Σενάριο 3 οι ερωτήσεις της Ελισάβετ ήταν επιβεβαιωτικές («Άρα...;»), σε συμφωνία με τις ορθές απαντήσεις του σεναρίου αναφοράς.	Σεν.2 Ερ.4· Σεν.3 Ερ.1–5
EE4	Όπου τέθηκε ερώτηση κατανόησης δόθηκε θετική απάντηση, οπότε ο μηχανισμός επανεξήγησης δεν ενεργοποιήθηκε.	Παραρτήματα Α–Γ

Πίνακας 5.21: Κύρια Ευρήματα

5.6 Προέλευση Δεδομένων και Έλεγχος Αριθμητικής Συμφωνίας

Οι τιμές των δεικτών ανά ερώτηση είναι αυτές που έδωσε το σύστημα στους διαλόγους των Παραρτημάτων Α–Γ. Οι πίνακες δείχνουν τη συνολική τιμή κάθε δείκτη και όχι τους 20 επιμέρους ελέγχους, οι οποίοι δεν καταγράφονται (Ενότητα 3.1.3).

Με τις τιμές των πινάκων υπολογίστηκαν ξανά οι μέσοι όροι και το Q. Οι μέσοι Q είναι 0,35, 0,36 και 1,00 και συμφωνούν με όσα ανέφερε το σύστημα.

5.7 Περιορισμοί Ανάλυσης

Οι μεθοδολογικοί περιορισμοί παρουσιάζονται στην Ενότητα 4.6 (Πίνακας 4.3). Για το παρόν κεφάλαιο επισημαίνεται επιπλέον ότι οι ερμηνείες στους πίνακες ανήκουν στον ερευνητή και στηρίζονται στα πλήρη εκφωνήματα, στις επιλογές και στις συνθήκες κάθε ερώτησης των Παραρτημάτων, όπως ίδιο βαρυτικό πεδίο ή τύπος ζυγού. Οι τιμές περιγράφουν τις συγκεκριμένες εκτελέσεις και όχι την επίδοση του συστήματος σε πραγματικό μαθητικό πληθυσμό.

Η ανάλυση ανέδειξε επίσης ότι, παρότι η διάσταση Consistency (C) έχει σχεδιαστεί ως διακριτή από τη Validity (V), στην εφαρμογή ορισμένων C checks μπορεί να εμφανιστεί εννοιολογική επικάλυψη μεταξύ συνέπειας και επιστημονικής ορθότητας. Το ζήτημα γίνεται εμφανές στο Σενάριο 1, όπου προέκυψε $C = 0,40$ χωρίς να εντοπίζεται ρητό ζεύγος ασύμβατων ισχυρισμών. Το εύρημα αυτό αναδεικνύει την ανάγκη σαφέστερης οριοθέτησης των C1–C5, ώστε η διάσταση C να αποτυπώνει με μεγαλύτερη ακρίβεια τη συνέπεια μεταξύ των θέσεων του μαθητή, ανεξάρτητα από την επιστημονική ορθότητά τους.

5.8 Σύνοψη Κεφαλαίου

Στα τρία συνθετικά σενάρια καταγράφονται διαφορετικές τιμές V, L, R, Clarity και C και ανατροφοδότηση με εννοιολογικές εξηγήσεις. Οι μέσοι Q υπολογίζονται σε 0,35, 0,36 και 1,00. Οι κύριες παρατηρήσεις (Ενότητα 5.5.3) αφορούν τη διαφοροποίηση των σεναρίων παρανοήσεων από το σενάριο αναφοράς στις συγκεκριμένες εκτελέσεις, τη διάκριση σαφήνειας και ορθότητας και την τήρηση της διαλογικής ακολουθίας.

6. Συμπεράσματα και Προτάσεις

Η εργασία εξέτασε το διαλογικό σύστημα Ph-D A.I. ως προς τρία σημεία: πώς ανταποκρίνεται σε προκαθορισμένες παρανοήσεις, τι ανατροφοδότηση δίνει και πώς εφαρμόζει το πλαίσιο VCLRCI-Q. Η διερεύνηση βασίστηκε σε τρία συνθετικά σενάρια με πέντε ερωτήσεις το καθένα, και τα συμπεράσματα αφορούν τη λειτουργία του πρωτοτύπου σε αυτές τις εκτελέσεις.

Το κεφάλαιο συνδέει τα ευρήματα με τα ερευνητικά ερωτήματα, παρουσιάζει τη συμβολή και τους περιορισμούς της εργασίας και προτείνει τα επόμενα βήματα για εμπειρική αξιολόγηση και τεχνολογική εξέλιξη.

6.1 Κύρια Συμπεράσματα

6.1.1 Ευρήματα από την Ανάλυση Σεναρίων

Στο Κεφάλαιο 5 οι επιστημονικά ορθές και οι εσφαλμένες αιτιολογήσεις πήραν σαφώς διαφορετικές τιμές δεικτών. Στα δύο σενάρια παρανοήσεων οι χαμηλοί μέσοι V συνυπήρχαν με υψηλότερες τιμές Clarity, δηλαδή το σύστημα ξεχώρισε τη σαφήνεια από την επιστημονική ορθότητα.

Οι μέσοι Q είναι 0,35, 0,36 και 1,00 και συμφωνούν με όσα ανέφερε το σύστημα (Ενότητα 5.6). Τα δύο σενάρια παρανοήσεων έχουν σχεδόν ίδιο Q και διαφέρουν καθαρά από το σενάριο αναφοράς.

Η ανατροφοδότηση κινήθηκε κυρίως στο επίπεδο έργου και διαδικασίας και ακολούθησε τη σχολική διάκριση μάζας και βάρους και τη σχέση $W = mg$. Η ακολουθία των τριών ρόλων τηρήθηκε σε όλες τις ερωτήσεις. Στο Σενάριο 2 η τιμή C αντιστοιχεί σε συγκεκριμένη αντίφαση μεταξύ των Ερ.1 και 5 (Ενότητα 5.3.3).

6.1.2 Σύνοψη Δεικτών

Δείκτης	Σεν.1	Σεν.2	Σεν.3	Παρατήρηση
V	0.26	0.24	1.00	Χαμηλότεροι μέσοι βαθμοί στα δύο σενάρια παρανοήσεων, υψηλός στο σενάριο αναφοράς.

Δείκτης	Σεν.1	Σεν.2	Σεν.3	Παρατήρηση
L	0.30	0.26	0.98	Χαμηλές τιμές στα σενάρια παρανοήσεων και υψηλή στο σενάριο αναφοράς.
R	0.56	0.82	1.00	Εκπαιδευτική καταλληλότητα για τη βαθμίδα. Υψηλότερη τιμή στο Σεν.2 από το Σεν.1.
Clarity	0.70	0.84	0.98	Η σαφής διατύπωση μπορεί να συνυπάρχει με επιστημονικά εσφαλμένο περιεχόμενο.
C	0.40	0.40	1.00	Συνέπεια μεταξύ των ισχυρισμών του μαθητή
Q	0.35	0.36	1.00	Διαφοροποίηση των σεναρίων παρανοήσεων από το σενάριο αναφοράς.

Πίνακας 6.1: Συνοπτική Συμπεριφορά Δεικτών στα Τρία Σενάρια

6.2 Απαντήσεις στα Ερευνητικά Ερωτήματα

EE1 Με ποιον τρόπο το διαλογικό σύστημα αναγνωρίζει και αποτυπώνει προκαθορισμένες παρανοήσεις στις αιτιολογήσεις των συνθετικών σεναρίων;

Το σύστημα εστίασε στη διάκριση μάζας και βάρους και στην εξάρτηση του βάρους από το βαρυτικό πεδίο. Η αιτιολόγηση έδωσε περισσότερο υλικό για εξέταση από την επιλογή μόνη της, και οι βαθμοί V, L, R και Clarity επέτρεψαν την περιγραφική ανάλυσή της. Στις συγκεκριμένες εκτελέσεις οι προκαθορισμένες παρανοήσεις αποτυπώθηκαν με χαμηλές τιμές V και L, ενώ οι ορθές απαντήσεις του σεναρίου αναφοράς πήραν υψηλές τιμές.

EE2 Ποια χαρακτηριστικά εμφανίζει η παραγόμενη ανατροφοδότηση και πώς συνδέονται με τις αρχές της εκπαιδευτικής ανατροφοδότησης των Hattie και Timperley (2007);

Η ανατροφοδότηση περιλαμβάνει αποσαφήνιση φυσικών εννοιών, εξήγηση σχέσεων και παραδείγματα. Τα χαρακτηριστικά αυτά αντιστοιχούν σε ανατροφοδότηση στο επίπεδο έργου και διαδικασίας (Hattie & Timperley, 2007· Shute, 2008).

EE3 Πώς αξιοποιούνται οι δείκτες Validity (V, εγκυρότητα), Consistency (C, συνέπεια), Logic (L, λογική), Regulation (R, εκπαιδευτική καταλληλότητα) και Clarity (Cl,

σαφήνεια) για την περιγραφική και συγκριτική ανάλυση των απαντήσεων και της αιτιολόγησης στα τρία συνθετικά σενάρια;

Ο V περιγράφει την επιστημονική ορθότητα, ο L τη λογική ακολουθία, ο R την εκπαιδευτική καταλληλότητα για τη βαθμίδα και η Clarity τη σαφήνεια. Ο C εξετάζει τη συνέπεια μεταξύ των αιτιολογήσεων και μπαίνει ως κοινή τιμή στο Q όλων των ερωτήσεων της συνεδρίας. Επειδή κάθε δείκτης δείχνει διαφορετική όψη της αιτιολόγησης, η σύγκριση των σεναρίων γίνεται πιο πλήρης. Για παράδειγμα, τα Σενάρια 1 και 2 είχαν το ίδιο ποσοστό ορθών επιλογών (1/5) και την ίδια τιμή C, αλλά διαφορετικές τιμές R και Clarity.

EE4 Πώς οργανώνεται η διαλογική αλληλεπίδραση μέσω των τριών παιδαγωγικών ρόλων και ποιες δυνατότητες αιτιολόγησης και αναστοχασμού δημιουργεί ο σχεδιασμός της;

Ο Κύριος Γιώργος δίνει τη διδακτική εξήγηση, η Ζωή κάνει σωκρατικές ερωτήσεις και η Ελισάβετ τον κριτικό έλεγχο. Η ακολουθία αυτή, μαζί με την απαίτηση αιτιολόγησης, δημιουργεί ευκαιρίες για επέκταση και αναστοχασμό, σε συνάφεια με τη διαλογική διδασκαλία (Alexander, 2008). Στους διαλόγους η ακολουθία των ρόλων τηρήθηκε σε όλες τις ερωτήσεις. Η Ζωή συνέδεσε σε μία περίπτωση τη συζήτηση με την καθημερινή έκφραση «ζυγίζω 50 kg», ενώ στο Σενάριο 3 οι ερωτήσεις της Ελισάβετ ήταν επιβεβαιωτικές, σε συμφωνία με τις σωστές απαντήσεις. Ο μηχανισμός επανεξήγησης δεν ενεργοποιήθηκε, καθώς όπου τέθηκε ερώτηση κατανόησης δόθηκε θετική απάντηση.

6.3 Θεωρητική Συμβολή της Εργασίας

6.3.1 Στη Διδακτική Φυσικών Επιστημών

Η συμβολή στη διδακτική των Φυσικών Επιστημών εστιάζει κυρίως στον σχεδιασμό. Η βιβλιογραφία για τις εναλλακτικές αντιλήψεις στη μάζα και στο βάρος αξιοποιήθηκε για τη διαμόρφωση συνθετικών εισόδων, οι οποίες συνδέθηκαν με την ανάλυση των αποκρίσεων του συστήματος. Με τον τρόπο αυτό, η εργασία συμβάλλει στον σχεδιασμό και στη διερευνητική εφαρμογή μιας οργανωμένης διαδικασίας διαλογικής αλληλεπίδρασης και αξιολόγησης.

6.3.2 Στην Αξιολόγηση και Ανατροφοδότηση

Η εργασία οργανώνει την επιστημονική ορθότητα, τη λογική ακολουθία, την εκπαιδευτική καταλληλότητα, τη σαφήνεια και τη συνέπεια ως διακριτές όψεις της αιτιολόγησης. Το

VCLRCI-Q χρησιμοποιείται ως πλαίσιο περιγραφής, όπως ορίζεται στο συγκεκριμένο prompt (τις οδηγίες προς το μοντέλο). Στην παρούσα μελέτη οι δείκτες αξιοποιούνται για την περιγραφική και συγκριτική ανάλυση των συγκεκριμένων εκτελέσεων, ενώ η διερεύνηση της σχέσης τους με τη μαθησιακή πρόοδο εντάσσεται σε επόμενο στάδιο εμπειρικής αξιολόγησης.

6.3.3 Στην Τεχνητή Νοημοσύνη στην Εκπαίδευση

Η εργασία συνδυάζει επιλεγμένες σχολικές σημειώσεις, δομημένες οδηγίες αξιολόγησης και τρεις παιδαγωγικούς ρόλους. Το μοντέλο λαμβάνει το κείμενο του επιλεγμένου κεφαλαίου από αρχείο TXT (απλό κείμενο), έως 20.000 χαρακτήρες, και οι οδηγίες του prompt το περιορίζουν σε αυτό το υλικό και του ζητούν να ελέγχει τις απαντήσεις του.

6.4 Παιδαγωγικές Προεκτάσεις

6.4.1 Για τον Εκπαιδευτικό

Η εφαρμογή μπορεί να δώσει στον εκπαιδευτικό χρήσιμο υλικό για:

1. αιτιολογήσεις που είναι σαφείς αλλά επιστημονικά λανθασμένες,
2. ισχυρισμούς που έρχονται σε αντίφαση μεταξύ τους,
3. παρανοήσεις που επανέρχονται στον διάλογο και
4. δηλώσεις άγνοιας ή αβεβαιότητας, που μπορούν να γίνουν αφετηρία συζήτησης στην τάξη.

Κάθε τιμή δείκτη ερμηνεύεται μαζί με την ερώτηση και τις συνθήκες της. Ο εκπαιδευτικός έχει τον κεντρικό ρόλο στην ερμηνεία και στις παιδαγωγικές αποφάσεις, και το σύστημα λειτουργεί υποστηρικτικά (Luckin et al., 2016).

6.4.2 Αρχές Σχεδιασμού Εκπαιδευτικών Συστημάτων Τεχνητής Νοημοσύνης

Από τον σχεδιασμό και τη δοκιμή του πρωτοτύπου προκύπτουν ορισμένες αρχές που αξίζει να αξιολογηθούν περαιτέρω:

- **Σύνδεση με επιλεγμένο υλικό:** η πηγή γνώσης είναι σαφώς ορισμένη. Μια μελλοντική επέκταση μπορεί να προσθέσει σημασιολογική αναζήτηση αποσπασμάτων.

- **Αιτιολόγηση πριν από την ανατροφοδότηση:** το σκεπτικό του μαθητή γίνεται ορατό και δίνει υλικό για διαμορφωτική αξιολόγηση (Chin & Osborne, 2010).
- **Ανατροφοδότηση πέρα από το σωστό/λάθος:** εξηγούνται οι έννοιες και ο συλλογισμός.
- **Παιδαγωγικοί ρόλοι:** οργανώνουν ερωτήσεις επέκτασης και κριτικού ελέγχου.

Η επίδραση αυτών των αρχών στη μάθηση, στον αναστοχασμό και στην εμπλοκή μπορεί να εξεταστεί με πραγματικούς μαθητές.

6.5 Περιορισμοί της Έρευνας

Η μελέτη αφορά τρία συνθετικά σενάρια της ενότητας «Μάζα και Βάρος». Κάθε σενάριο εκτελέστηκε μία φορά, με την προεπιλεγμένη ρύθμιση temperature (τυχειότητα των απαντήσεων) του μοντέλου, και τους βαθμούς έδωσε το γλωσσικό μοντέλο. Οι υπόλοιποι μεθοδολογικοί περιορισμοί συνοψίζονται στον Πίνακα 4.3. Η σταθερότητα των βαθμών σε επαναλαμβανόμενες εκτελέσεις, η ανεξάρτητη βαθμολόγηση από ειδικούς, η εφαρμογή με πραγματικούς μαθητές και η μαθησιακή αποτελεσματικότητα του συστήματος αποτελούν το επόμενο ερευνητικό στάδιο (Ενότητα 6.6).

6.6 Προοπτικές Περαιτέρω Διερεύνησης

6.6.1 Εμπειρική Αξιολόγηση

Το επόμενο βήμα είναι μια μελέτη σε πραγματικές σχολικές συνθήκες, με σαφές ερευνητικό πρωτόκολλο. Μια πιθανή προσέγγιση περιλαμβάνει:

1. **Συμμετέχοντες:** μαθητές Γυμνασίου από μία ή περισσότερες σχολικές μονάδες, με καθορισμένα κριτήρια επιλογής και τις απαραίτητες εγκρίσεις.
2. **Σχεδιασμός:** αρχική και τελική αξιολόγηση με ανεξάρτητες ερωτήσεις, χρήση του συστήματος ανάμεσα στις δύο και, αν είναι εφικτό, μεταγενέστερη μέτρηση για το πόσο διατηρήθηκε η γνώση.
3. **Ομάδα σύγκρισης:** η ίδια ενότητα με άλλη διδακτική προσέγγιση, στον ίδιο χρόνο και με τα ίδια κριτήρια αξιολόγησης.

4. **Δεδομένα:** αποτελέσματα των ανεξάρτητων δοκιμασιών, πλήρεις διάλογοι, αναλυτικές εγγραφές και στοιχεία για την εμπειρία των μαθητών. Οι τιμές VCLRCI-Q χρησιμοποιούνται περιγραφικά, δίπλα στην ανεξάρτητη μέτρηση της μάθησης.

Για τη συλλογή πραγματικών μαθητικών δεδομένων χρειάζεται πρωτόκολλο ενημέρωσης, συναίνεσης, πρόσβασης, διατήρησης και διαγραφής των δεδομένων (Ενότητα 4.7). Στην τρέχουσα μορφή της η εφαρμογή χρησιμοποιεί ελεύθερη εισαγωγή username (ονόματος χρήστη) και αυτόματη αποθήκευση, οπότε το πρωτόκολλο αυτό θα προστεθεί πριν από τη χρήση της με μαθητές.

6.6.2 Επικύρωση Δεικτών

Για να επικυρωθεί το VCLRCI-Q χρειάζονται αυθεντικές μαθητικές απαντήσεις και προκαθορισμένα κριτήρια. Τρεις διαδικασίες μπορούν να το εξετάσουν:

1. **Σύγκριση με ειδικούς:** ανεξάρτητοι ειδικοί βαθμολογούν τις ίδιες αιτιολογήσεις με την ίδια ρουμπρίκα, χωρίς να βλέπουν τις τιμές του μοντέλου. Συγκρίνονται τόσο οι κρίσεις TRUE (ναι), PARTIAL (μερικώς) και FALSE (όχι) κάθε ελέγχου όσο και οι συνολικοί βαθμοί, και οι διαφορές εξετάζονται μαζί με το κείμενο της απάντησης. Οι ειδικοί ελέγχουν επίσης αν κάθε κριτήριο μετρά πράγματι αυτό που ορίζει. Ειδικά για τη διάσταση C, η διαδικασία αυτή μπορεί να υποστηρίξει σαφέστερη οριοθέτηση των C1–C5 από τη διάσταση Validity (V), ώστε να περιοριστεί πιθανή επικάλυψη μεταξύ συνέπειας και επιστημονικής ορθότητας. Συμπληρωματικά, μπορεί να δοθεί δεύτερη κρίση από άλλο LLM (μεγάλο γλωσσικό μοντέλο).
2. **Επαναλαμβανόμενες εκτελέσεις:** με τις ίδιες απαντήσεις, το ίδιο prompt, τις ίδιες σημειώσεις και καταγεγραμμένη έκδοση μοντέλου, εξετάζεται πόσο σταθεροί είναι οι βαθμοί. Η σταθερότητα αυτή είναι διαφορετικό ζήτημα από τον δείκτη C, που αφορά τη συνέπεια του μαθητή.
3. **Σύνδεση με τη μάθηση:** οι βαθμοί συγκρίνονται με ανεξάρτητες δοκιμασίες γνώσεων.

Στην παρούσα μελέτη το Q λειτουργεί ως συνδυαστική περιγραφή των δεικτών και όχι ως μέτρηση μιας ενιαίας ικανότητας. Μελλοντικά μπορεί να εξεταστεί η αιτιολόγηση των βαρών και το πώς αλλάζουν τα αποτελέσματα με διαφορετική στάθμιση. Οι διαδικασίες αυτές θα δείξουν σε ποιες συνθήκες οι δείκτες είναι αξιόπιστοι και χρήσιμοι.

6.6.3 Επέκταση Περιεχομένου

Το σύστημα φορτώνει τις σημειώσεις ανά βιβλίο και κεφάλαιο, οπότε μπορεί να επεκταθεί σε νέο περιεχόμενο με την προσθήκη νέων αρχείων σημειώσεων. Η προσέγγιση μπορεί να εφαρμοστεί και να αξιολογηθεί:

1. σε άλλες ενότητες της Φυσικής, όπως οι Δυνάμεις, ο Ηλεκτρισμός, τα Κύματα και η Οπτική,
2. σε άλλες τάξεις του Γυμνασίου και του Λυκείου και
3. σε άλλα μαθήματα, όπως η Χημεία (π.χ. γραμμομοριακές έννοιες) και η Βιολογία (π.χ. κυτταρική διαίρεση).

6.6.4 Τεχνολογικές Βελτιώσεις

Οι τεχνολογικές βελτιώσεις αφορούν κυρίως την αποθήκευση των δεδομένων, τον αριθμητικό έλεγχο και τις νέες λειτουργίες για τον εκπαιδευτικό και τον μαθητή.

1. **Αποθήκευση δεδομένων:** σήμερα οι αναλυτικές εγγραφές κάθε ερώτησης (Q_RECORD) μένουν στον φυλλομετρητή και ο διακομιστής αποθηκεύει τη σύνοψη του κουίζ (QUIZ_SUMMARY). Το πεδίο questions (ερωτήσεις) μένει σκόπιμα κενό, ώστε να μη μεγαλώνει ο όγκος των αρχείων· σε μια ερευνητική εκδοχή μπορεί να συμπληρώνεται, ώστε να αποθηκεύεται ολόκληρο το κουίζ, μαζί με την πλήρη διατύπωση κάθε ερώτησης και τις επιλογές της.
2. **Αριθμητικός έλεγχος:** ο υπολογισμός των μέσων όρων και του Q μπορεί να γίνεται και στον διακομιστή, ώστε να επαληθεύονται αυτόματα οι τιμές του γλωσσικού μοντέλου.
3. **Πίνακας εκπαιδευτικού:** ένας πίνακας παρακολούθησης μπορεί να αξιοποιεί τις αποθηκευμένες συνόψεις, ξεχωρίζοντας τους βαθμούς του μοντέλου από τα ανεξάρτητα στοιχεία μάθησης.
4. **Προσαρμογή δυσκολίας:** το σύστημα ήδη στέλνει στο μοντέλο έως τρεις προηγούμενες συνόψεις του χρήστη. Αυτό δίνει τη βάση για προσαρμογή της δυσκολίας με σαφή κανόνα, π.χ. ανά ενότητα και τάξη.
5. **Νέες μορφές:** εικόνες, προσομοιώσεις και φωνητική διεπαφή.

Κάθε νέα λειτουργία χρειάζεται δοκιμή με χρήστες και ειδικούς, ως προς τη χρησιμότητα και την παιδαγωγική της αξία.

6.7 Συνοπτική Αποτίμηση της Εργασίας

Η εργασία υλοποίησε και εξέτασε ένα πρωτότυπο που συνδυάζει σχολικές σημειώσεις, τρεις διαλογικούς παιδαγωγικούς ρόλους, απαίτηση αιτιολόγησης και το πλαίσιο VCLRCI-Q. Οι τρεις συνθετικές εκτελέσεις έδωσαν υλικό για την ανάλυση των εξηγήσεων και των βαθμών του συστήματος.

Στις εκτελέσεις αυτές οι δείκτες ξεχώρισαν τη σαφήνεια από την επιστημονική ορθότητα και διέκριναν καθαρά τα σενάρια παρανοήσεων από το σενάριο αναφοράς. Η συμβολή της εργασίας βρίσκεται στον σχεδιασμό αυτής της διαδικασίας και στην πρώτη, διερευνητική εφαρμογή της. Τα επόμενα βήματα είναι η αξιολόγηση των δεικτών από ανεξάρτητους ειδικούς και η εφαρμογή του συστήματος με πραγματικούς μαθητές, ώστε να εξεταστεί η συμβολή του στη μάθηση.

Βιβλιογραφία

- Alexander, R. J. (2008). *Towards dialogic teaching: Rethinking classroom talk* (4th ed.). Dialogos.
- Anthropic. (n.d.). *Prompt caching*. Ανακτήθηκε από
<https://platform.claude.com/docs/en/build-with-claude/prompt-caching>
- ASSISTments. (n.d.). *ASSISTments for teachers*. Ανακτήθηκε από
<https://www.assistments.org/product/teacher>
- Chin, C., & Osborne, J. (2010). Supporting argumentation through students' questions: Case studies in science classrooms. *Journal of the Learning Sciences*, 19(2), 230–284.
<https://doi.org/10.1080/10508400903530036>
- Driver, R., Guesne, E., & Tiberghien, A. (Eds.). (1985). *Children's ideas in science*. Open University Press.
- Galili, I. (1993). Weight and gravity: Teachers' ambiguity and students' confusion about the concepts. *International Journal of Science Education*, 15(2), 149–162.
<https://doi.org/10.1080/0950069930150203>
- Graesser, A. C., Chipman, P., Haynes, B. C., & Olney, A. (2005). AutoTutor: An intelligent tutoring system with mixed-initiative dialogue. *IEEE Transactions on Education*, 48(4), 612–618. <https://doi.org/10.1109/TE.2005.856149>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., . . . Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274.
<https://doi.org/10.1016/j.lindif.2023.102274>
- Khan Academy. (n.d.). *Khanmigo for learners*. Ανακτήθηκε από
<https://www.khanmigo.ai/learners>

- Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson.
- Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, 66(2), 211–227. <https://doi.org/10.1002/sce.3730660207>
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153–189. <https://doi.org/10.3102/0034654307313795>
- Viennot, L. (1979). Spontaneous reasoning in elementary dynamics. *European Journal of Science Education*, 1(2), 205–221. <https://doi.org/10.1080/0140528790010209>
- Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE.
- Zheng, L., Chiang, W.-L., Sheng, Y., Zhuang, S., Wu, Z., Zhuang, Y., Lin, Z., Li, Z., Li, D., Xing, E. P., Zhang, H., Gonzalez, J. E., & Stoica, I. (2023). Judging LLM-as-a-judge with MT-Bench and Chatbot Arena. *Advances in Neural Information Processing Systems*, 36.

Παραρτήματα — Διάλογοι των Συνθετικών Σεναρίων

Τα Παραρτήματα Α–Γ περιέχουν τους διαλόγους των τριών συνθετικών σεναρίων που αναλύονται στο Κεφάλαιο 5. Τα κείμενα των ερωτήσεων, των απαντήσεων, της ανατροφοδότησης και του σωκρατικού διαλόγου αντιγράφηκαν χειροκίνητα από τη διεπαφή της εφαρμογής. Οι πίνακες βαθμών κάθε ερώτησης αποδίδονται συνοπτικά σε μία γραμμή και οι επισημάνσεις «ΣΥΣΤΗΜΑ», «ΜΑΘΗΤΗΣ» και «ΑΝΑΤΡΟΦΟΔΟΤΗΣΗ» προστέθηκαν για αναγνωσιμότητα. Η «Σύνοψη ερευνητή» στο τέλος κάθε σεναρίου καταγράφει τις συγκεντρωτικές τιμές και τους ελέγχους συνέπειας C1–C5, όπως τους ανέφερε το σύστημα στην τελική αξιολόγηση, καθώς και τις επανυπολογισμένες τιμές που χρησιμοποιούνται στην ανάλυση του Κεφαλαίου 5. Οι πρότυπες δομές Q_RECORD και QUIZ_SUMMARY παρατίθενται στο Παράρτημα Δ. Το περιεχόμενο των διαλόγων διατηρείται όπως εμφανίστηκε στην εφαρμογή, συμπεριλαμβανομένων τυχόν γλωσσικών ή επιστημονικών αστοχιών της παραγόμενης ανατροφοδότησης, ώστε να μην αλλοιώνεται το υλικό που αναλύεται στην εργασία.

Παράρτημα Α : Σενάριο 1 (Μαθητής_A)

Συνθετικό προφίλ: μεταβολή της μάζας στη Σελήνη και σύγκριση μάζας και βάρους (M_1 , M_2)

Username: mathitis_A | Τάξη: Β' Γυμνασίου | Ερωτήσεις: 5

ΣΥΣΤΗΜΑ: Καλώς ήρθες στο Physics Quiz! 📖

Βλέπω ότι θα εξεταστείς στο κεφάλαιο που επέλεξες.

Χρειάζομαι ακόμα:

□ Ποια τάξη είσαι;

▢ Πόσες ερωτήσεις θέλεις; (Προτείνω 5-10)

ΜΑΘΗΤΗΣ: Β' Γυμνασίου 5

ΣΥΣΤΗΜΑ: Τέλεια! 🎯 Είσαι Β' Γυμνασίου και θα λύσεις 5 ερωτήσεις στο κεφάλαιο

Μάζα και Βάρος.

Ετοιμάζω τις ερωτήσεις... ⌚

?Ερώτηση 1/5

ΣΥΣΤΗΜΑ: Ποιο από τα παρακάτω είναι χαρακτηριστικό της μάζας;

- A. Αλλάζει ανάλογα με το ύψος από την επιφάνεια της Γης
 - B. Εκφράζει την ποσότητα ύλης που περιέχει το σώμα
 - C. Μετρείται σε Newton (N)
 - D. Παίρνει διαφορετικές τιμές σε διαφορετικούς πλανήτες
- Ποια είναι η σωστή απάντηση; (Απάντησε A, B, C ή D)

ΜΑΘΗΤΗΣ: δ

ΣΥΣΤΗΜΑ: Διάλεξες D. Εξήγησέ μου γιατί — τι σκέφτηκες;

ΜΑΘΗΤΗΣ: Γιατί στη σελήνη είμαστε ελαφρύτεροι, οπότε μειώνεται η μάζα μας.

ΑΝΑΤΡΟΦΟΔΟΤΗΣΗ: ✗ Σωστή απάντηση: B

  Κύριος Γιώργος: Κατάλαβα τη σκέψη σου — πράγματι στη Σελήνη νιώθουμε ελαφρύτεροι! Αλλά αυτό συμβαίνει γιατί μειώνεται το βάρος, όχι η μάζα. Η μάζα είναι η ποσότητα ύλης και παραμένει ΠΑΝΤΑ ίδια — είτε στη Γη, είτε στη Σελήνη, είτε στο διάστημα! Ένας αστροναύτης που ζυγίζει 60 κιλά στη Γη, έχει ακριβώς 60 κιλά μάζα και στη Σελήνη. Αλλά το βάρος του γίνεται 6 φορές μικρότερο, γιατί στη Σελήνη το g είναι 6 φορές μικρότερο.

 V=0.1 | L=0.2 | R=0.5 | Clarity=0.7 | C=

 Ζωή: Άρα στη Σελήνη αν πάμε με ένα ζυγό και ένα δυναμόμετρο, θα δείξουν διαφορετικά νούμερα για το ίδιο σώμα;

  Κύριος Γιώργος: Ακριβώς! Ο ζυγός (που μετράει μάζα) θα δείξει το ίδιο. Το δυναμόμετρο (που μετράει βάρος) θα δείξει 6 φορές μικρότερη τιμή.

 Ελισάβετ: Αν κάποιος ήταν σε ένα διαστημόπλοιο μακριά από κάθε πλανήτη, θα είχε μάζα;

  Κύριος Γιώργος: Ναι, φυσικά! Η μάζα του σώματος θα ήταν η ίδια — το υλικό δεν αλλάζει. Αλλά το βάρος θα ήταν (σχεδόν) μηδέν, γιατί δεν υπάρχει βαρυτική έλξη.

  Κύριος Γιώργος: Το κατάλαβες; Θες να το ξαναδούμε ή πάμε στην επόμενη ερώτηση;

ΜΑΘΗΤΗΣ: Ναι, κατάλαβα! Επόμενη →

?Ερώτηση 2/5

ΣΥΣΤΗΜΑ: Ένας μαθητής θέλει να μετρήσει τη μάζα ενός άγνωστου αντικειμένου. Ποιο όργανο πρέπει να χρησιμοποιήσει;

- A. Δυναμόμετρο
- B. Ζυγός σύγκρισης
- C. Θερμόμετρο
- D. Χάρακας

ΜΑΘΗΤΗΣ: α

ΣΥΣΤΗΜΑ: Διάλεξες Α. Εξήγησέ μου γιατί — τι σκέφτηκες;

ΜΑΘΗΤΗΣ: Γιατί όσο πιο βαρύ είναι ένα σώμα τόσο πιο πολύ δύναμη ασκεί.

ΑΝΑΤΡΟΦΟΔΟΤΗΣΗ: ✗ Σωστή απάντηση: B

  Κύριος Γιώργος: Έχεις δίκιο ότι υπάρχει σχέση μεταξύ μάζας και δύναμης, αλλά δεν είναι αυτή η σωστή σχέση εδώ. Δυναμόμετρο: Μετράει τη δύναμη (το βάρος) — και το βάρος αλλάζει ανάλογα με τον τόπο! Ζυγός σύγκρισης: Μετράει τη μάζα συγκρίνοντας το αντικείμενο με πρότυπες μάζες. Δουλεύει παντού — στη Γη, στη Σελήνη και στο διάστημα — γιατί συγκρίνει ποσότητες ύλης, όχι δυνάμεις.

 V=0.2 | L=0.3 | R=0.5 | Clarity=0.8 | C=

 Ζωή: Αν ένας ζυγός είναι σπασμένος και δεν ισορροπεί καλά, μπορώ ακόμα να μετρήσω μάζα;

  Κύριος Γιώργος: Καλή ερώτηση! Ναι, αν αγνοήσεις το σφάλμα. Αλλά αν θέλεις ακριβή αποτέλεσμα, χρειάζεσαι έναν σωστό ζυγό που να ισορροπεί σωστά.

 Ελισάβετ: Γιατί λοιπόν υπάρχουν ηλεκτρονικοί ζυγοί που δείχνουν αριθμούς σε κιλά και όχι δυναμόμετρα;

  Κύριος Γιώργος: Ο ηλεκτρονικός ζυγός στην πραγματικότητα μετράει το βάρος με έναν αισθητήρα και στη συνέχεια διαιρεί με το g για να σου δείξει τη μάζα. Αυτό λειτουργεί τέλεια στη Γη, αλλά σε άλλο πλανήτη θα δείξει λάθος αποτέλεσμα!

  Κύριος Γιώργος: Το κατάλαβες; Θες να το ξαναδούμε ή πάμε στην επόμενη ερώτηση;

ΜΑΘΗΤΗΣ: Ναι, κατάλαβα! Επόμενη →

?Ερώτηση 3/5

ΣΥΣΤΗΜΑ: Η επιμήκυνση ενός ελατηρίου είναι ανάλογη της μάζας που κρεμάμε σε αυτό, όταν βρισκόμαστε στον ίδιο τόπο. Ποια είναι η σωστή μαθηματική σχέση;

A. $x = m^2/k$

B. $x = k/m$

C. $x = mg/k$

D. $x = k \cdot m$

ΜΑΘΗΤΗΣ: A

ΣΥΣΤΗΜΑ: Διάλεξες A. Εξήγησέ μου γιατί — τι σκέφτηκες;

ΜΑΘΗΤΗΣ: Στην τύχη.

ΑΝΑΤΡΟΦΟΔΟΤΗΣΗ: ✗ Σωστή απάντηση: C

  Κύριος Γιώργος: Όταν κρεμάμε μια μάζα στο ελατήριο, το βάρος της ασκεί δύναμη προς τα κάτω. Ελατήριο: $mg = kx \rightarrow x = mg/k$. Όσο μεγαλύτερη η μάζα $\rightarrow$ μεγαλύτερη επιμήκυνση. Αν αλλάξεις πλανήτη (αλλάζει g), η ίδια μάζα θα τεντώσει λιγότερο το ελατήριο.

 V=0.0 | L=0.0 | R=0.5 | Clarity=0.5 | C=

 Ζωή: Αν διπλασιάσουμε τη μάζα, η επιμήκυνση θα διπλασιαστεί;

  Κύριος Γιώργος: Ναι, ακριβώς! Γιατί στον τύπο $x = mg/k$, η επιμήκυνση είναι ανάλογη της μάζας. Διπλάσια m $\rightarrow$ διπλάσιο x.

 Ελισάβετ: Τι θα συνέβαινε αν χρησιμοποιούσαμε το ίδιο ελατήριο με την ίδια μάζα αλλά στη Σελήνη;

  Κύριος Γιώργος: Στη Σελήνη το g είναι 6 φορές μικρότερο, οπότε η επιμήκυνση θα ήταν 6 φορές μικρότερη. Η μάζα δεν αλλάζει, αλλά η επιμήκυνση αλλάζει γιατί αλλάζει το g.

  Κύριος Γιώργος: Το κατάλαβες; Θες να το ξαναδούμε ή πάμε στην επόμενη ερώτηση;

ΜΑΘΗΤΗΣ: Ναι, κατάλαβα! Επόμενη $\rightarrow$

?Ερώτηση 4/5

ΣΥΣΤΗΜΑ: Ένα σώμα έχει μάζα 4 kg. Ποιο είναι το βάρος του στη Γη; ($g=10 \text{ m/s}^2$)

A. 4 N

- B. 20 N
C. 40 N
D. 0,4 N

ΜΑΘΗΤΗΣ: α

ΣΥΣΤΗΜΑ: Διάλεξες Α. Εξήγησέ μου γιατί — τι σκέφτηκες;

ΜΑΘΗΤΗΣ: Γιατί μάζα και Βάρος είναι το ίδιο.

ΑΝΑΤΡΟΦΟΛΟΤΗΣΗ: ✗ Σωστή απάντηση: C

  Κύριος Γιώργος: Μάζα και βάρος είναι εντελώς διαφορετικά! Μάζα (kg) = ποσότητα ύλης. Βάρος (N) = δύναμη βαρύτητας. $W = mg = 4 \times 10 = 40 \text{ N}$. Ένας μαθητής 40kg: μάζα=40kg παντού, βάρος στη Γη=400N, βάρος στη Σελήνη≈67N.

 V=0.0 | L=0.1 | R=0.5 | Clarity=0.6 | C=

 Ζωή: Αν ένα σώμα έχει μάζα 2 kg, πόσο θα είναι το βάρος του;

  Κύριος Γιώργος: Με $g=10$, το βάρος θα είναι $W=2 \times 10=20\text{N}$. Βλέπεις; Η μάζα 2kg δίνει βάρος 20N, όχι 2N!

 Ελισάβετ: Γιατί λοιπόν στη Σελήνη ένας αστροναύτης με μάζα 60kg δεν ζυγίζει 60N;

  Κύριος Γιώργος: Επειδή στη Σελήνη το g είναι μικρότερο! Αν $g_{\text{σελ}}=1.6$, τότε $W=60 \times 1.6=96\text{N}$. Η μάζα είναι πάντα 60kg, αλλά το βάρος αλλάζει με το g .

  Κύριος Γιώργος: Το κατάλαβες τώρα; Θες να το ξαναδούμε ή πάμε στην επόμενη ερώτηση;

ΜΑΘΗΤΗΣ: Ναι, κατάλαβα! Επόμενη →

?Ερώτηση 5/5

ΣΥΣΤΗΜΑ: Δύο σώματα έχουν την ίδια μάζα και βρίσκονται στη Γη. Ποια από τις παρακάτω δηλώσεις είναι σωστή;

- A. Το ένα σώμα έχει μεγαλύτερη μάζα από το άλλο
B. Τα δύο σώματα έχουν ίδιο βάρος
C. Το βάρος τους εξαρτάται από το σχήμα τους
D. Η μάζα τους θα διαφέρει αν τα μεταφέρουμε στη Σελήνη

ΜΑΘΗΤΗΣ: β

ΣΥΣΤΗΜΑ: Διάλεξες Β. Εξήγησέ μου γιατί — τι σκέφτηκες;

ΜΑΘΗΤΗΣ: Αφού έχουν ίδια μάζα, θα έχουν και ίδιο βάρος.

ΑΝΑΤΡΟΦΟΔΟΤΗΣΗ:  Σωστά! Απάντηση Β είναι η σωστή!

  Κύριος Γιώργος: Εξαιρετικό συλλογισμό! $W=mg$. Ίδια m + ίδιο g (ίδιος τόπος) = ίδιο βάρος! Αν τα μεταφέρουμε στη Σελήνη, η μάζα παραμένει ίδια (0.2kg), αλλά το βάρος θα γίνει 6 φορές μικρότερο.

 $V=1.0$ | $L=0.9$ | $R=0.8$ | Clarity=0.9 | C=

 Ζωή: Αν πάμε και τα δύο σώματα στη Σελήνη, θα έχουν πάλι ίδιο βάρος;

  Κύριος Γιώργος: Ναι! Θα έχουν πάλι ίδιο βάρος — μικρότερο από τη Γη, αλλά ίδιο μεταξύ τους, γιατί ίδια μάζα + ίδιο g σελ.

 Ελισάβετ: Τι θα γίνονταν αν ένα από τα δύο σώματα ήταν πολύ πιο κοντά στο κέντρο της Γης — θα άλλαζε το βάρος;

  Κύριος Γιώργος: Ναι, θα άλλαζε λίγο. Όσο πιο κοντά στο κέντρο, τόσο μεγαλύτερο το g . Αλλά στις καθημερινές αποστάσεις, η διαφορά είναι αμελητέα.

Σύνοψη ερευνητή — Σενάριο 1

Τιμές που ανέφερε το σύστημα: avg_V=0.26 | C=0.40 | avg_L=0.30 | avg_R=0.56 |
avg_Clarify=0.70 | Q=0.35

Έλεγχοι συνέπειας: C1=PARTIAL(0.1) | C2=PARTIAL(0.1) | C3=FALSE(0.0) |
C4=FALSE(0.0) | C5=TRUE(0.2)

Σωστές απαντήσεις: 1/5

Παράρτημα Β — Σενάριο 2 (Μαθητής_B)

Συνθετικό προφίλ: το βάρος ως σταθερή ιδιότητα του σώματος (M3)

Username: mathitis_B | Τάξη: Β' Γυμνασίου | Ερωτήσεις: 5

ΣΥΣΤΗΜΑ: Καλώς ήρθες στο Physics Quiz! 📖

Βλέπω ότι θα εξεταστείς στο κεφάλαιο "Μάζα και Βάρος" (Γυμνάσιο).

Χρειάζομαι:

□ Ποια τάξη είσαι;

□ Πόσες ερωτήσεις θέλεις;

ΜΑΘΗΤΗΣ: Β' Γυμνασίου 5

ΣΥΣΤΗΜΑ: Τέλεια! 🎯 Είσαι Β' Γυμνασίου και θα κάνουμε 5 ερωτήσεις στο κεφάλαιο "Μάζα και Βάρος".

Ετοιμάζω τις ερωτήσεις... ⌚

? Ερώτηση 1/5

ΣΥΣΤΗΜΑ: Ποιο από τα παρακάτω είναι σταθερό για ένα σώμα, ανεξάρτητα από τον τόπο που βρίσκεται;

- A. Το βάρος του
- B. Η μάζα του
- C. Η δύναμη που του ασκεί η Γη
- D. Η επιμήκυνση ενός ελατηρίου στο οποίο κρέμεται

ΜΑΘΗΤΗΣ: Α

ΣΥΣΤΗΜΑ: Διάλεξες Α. Εξήγησέ μου γιατί — τι σκέφτηκες;

ΜΑΘΗΤΗΣ: Το βάρος δεν αλλάζει, είναι το ίδιο παντού.

ΑΝΑΤΡΟΦΟΛΟΤΗΣΗ: ✗ Σωστή απάντηση: Β

🧐 📖 Κύριος Γιώργος: Καταλαβαίνω τη σκέψη σου, αλλά εδώ είναι ακριβώς το αντίθετο!

Το βάρος αλλάζει ανάλογα με τον τόπο. $W = mg$, όπου g δεν είναι σταθερό παντού. Στη Σελήνη $g \approx 1.6 \text{ m/s}^2$. Αστροναύτης 60kg: $W_{\text{γη}} = 600\text{N}$, $W_{\text{σελ}} = 96\text{N}$. Η μάζα παραμένει πάντα 60kg.

 V=0.0 | L=0.1 | R=0.8 | Clarity=0.9 | C=

☺ Ζωή: Και αν πάμε σε ένα ψηλότερο βουνό στη Γη — αλλάζει και εκεί το βάρος μας;

☺ 🏠 Κύριος Γιώργος: Ναι! Όσο ανεβαίνουμε ψηλότερα, το g μειώνεται λίγο, οπότε το βάρος μας μειώνεται ελάχιστα.

☺ Ελισάβετ: Αλλά γιατί τότε ο ζυγός του σπιτιού δείχνει ίδιο αποτέλεσμα όταν ανεβούμε βουνό;

☺ 🏠 Κύριος Γιώργος: Επειδή η μεταβολή του g είναι πολύ μικρή στα ύψη που συναντάμε καθημερινά. Πρέπει να πάμε πολύ πιο ψηλά ή σε άλλο πλανήτη για να δούμε σημαντική διαφορά.

☺ 🏠 Κύριος Γιώργος: Το κατάλαβες; Θες να το ξαναδούμε ή πάμε στην επόμενη ερώτηση;

ΜΑΘΗΤΗΣ: Ναι, κατάλαβα! Επόμενη →

? Ερώτηση 2/5

ΣΥΣΤΗΜΑ: Ένας μαθητής χρησιμοποιεί έναν ηλεκτρονικό ζυγό στη Γη και μετράει τη μάζα ενός σώματος: 2 kg. Αν πάει στη Σελήνη με το ίδιο σώμα και το ζυγό, τι θα δείξει ο ζυγός;

- A. 2 kg (το ίδιο)
- B. Περίπου 0,33 kg (6 φορές μικρότερο)
- C. Περίπου 12 kg (6 φορές μεγαλύτερο)
- D. 0 kg (δεν έχει βάρος στη Σελήνη)

ΜΑΘΗΤΗΣ: α

ΣΥΣΤΗΜΑ: Διάλεξες Α. Εξήγησέ μου γιατί — τι σκέφτηκες;

ΜΑΘΗΤΗΣ: Η μάζα είναι 2kg γιατί το βάρος δεν αλλάζει, είναι χαρακτηριστικό του σώματος.

ΑΝΑΤΡΟΦΟΔΟΤΗΣΗ: ✗ Σωστή απάντηση: B

☺ 🏠 Κύριος Γιώργος: Ηλεκτρονικός ζυγός δεν μετράει τη μάζα απευθείας. Δουλεύει μέσω βάρους: $m = W/g$. Στη Γη ($g=10$): $W=20N \rightarrow$ ζυγός δείχνει 2kg. Στη Σελήνη ($g \approx 1.6$):

$W=3.2\text{N} \rightarrow$ ζυγός υπολογίζει $m=3.2/10=0.32\text{kg}\approx 0.33\text{kg}$. Ζυγός σύγκρισης θα έδειχνε σωστά 2kg παντού.

 $V=0.2 \mid L=0.2 \mid R=0.8 \mid \text{Clarity}=0.8 \mid C=\text{⌚}$

 Ζωή: Άρα ο ηλεκτρονικός ζυγός είναι «ηλίθιος» — δεν ξέρει ότι είναι στη Σελήνη;

  Κύριος Γιώργος: Ακριβώς! Είναι προγραμματισμένος για τη Γη. Νομίζει ότι ο μικρός αριθμός που μετράει σημαίνει μικρή μάζα.

 Ελισάβετ: Τότε πώς θα μπορούσε ένας αστροναύτης να μετρήσει σωστά τη μάζα ενός αντικειμένου στη Σελήνη;

  Κύριος Γιώργος: Θα έπρεπε να χρησιμοποιήσει ένα ζυγό σύγκρισης με γνωστές μάζες αναφοράς — όχι ηλεκτρονικό!

  Κύριος Γιώργος: Πάμε στην επόμενη ερώτηση! 

ΜΑΘΗΤΗΣ: επόμενη ερώτηση

?Ερώτηση 3/5

ΣΥΣΤΗΜΑ: Ένα ελατήριο κρέμεται κατακόρυφα. Όταν προσθέτουμε μάζες στο άκρο του, η επιμήκυνσή του αυξάνεται. Τι είδους σχέση υπάρχει μεταξύ μάζας m και επιμήκυνσης x στον ίδιο τόπο;

- A. Η επιμήκυνση είναι ανάλογη της τετραγώνου της μάζας: $x \sim m^2$
- B. Η επιμήκυνση είναι αντιστρόφως ανάλογη της μάζας: $x \sim 1/m$
- C. Η επιμήκυνση είναι ευθέως ανάλογη της μάζας: $x \sim m$
- D. Δεν υπάρχει σχέση — η επιμήκυνση είναι τυχαία

ΜΑΘΗΤΗΣ: γ

ΣΥΣΤΗΜΑ: Διάλεξες Γ (δηλαδή C). Εξήγησέ μου γιατί — τι σκέφτηκες;

ΜΑΘΗΤΗΣ: Όσο βαρύτερο το σώμα τόσο τεντώνει το ελατήριο, άρα το βάρος και η επιμήκυνση είναι ίδια πράγμα.

ΑΝΑΤΡΟΦΟΔΟΤΗΣΗ:  Σωστά! Απάντηση C είναι η σωστή!

  Κύριος Γιώργος: Πολύ καλή σκέψη! Ισορροπία: $mg = kx \rightarrow x = mg/k$. Σταθερό g και $k \rightarrow x$ ανάλογο m . Διπλάσια μάζα $\rightarrow$ διπλάσια επιμήκυνση! Στη Σελήνη g μικρότερο $\rightarrow$ η ίδια μάζα τεντώνει λιγότερο το ελατήριο.

 V=1.0 | L=0.9 | R=1.0 | Clarity=0.9 | C= Ζωή: Αν ένα ελατήριο τεντωθεί 10 cm με 2 kg, πόσο θα τεντωθεί με 6 kg;  Κύριος Γιώργος: Αφού η σχέση είναι ανάλογη, τριπλάσια μάζα → τριπλάσια επιμήκυνση → 30 cm! Ελισάβετ: Και αν το ελατήριο έχει διαφορετική σταθερά k — πιο σκληρό ή πιο μαλακό;  Κύριος Γιώργος: Η σχέση $x=mg/k$ εξακολουθεί να ισχύει! Σκληρότερο ελατήριο (μεγαλύτερο k) → μικρότερη επιμήκυνση για ίδια μάζα.  Κύριος Γιώργος: Πολύ καλά! Πάμε στην επόμενη ερώτηση! 

?Ερώτηση 4/5

ΣΥΣΤΗΜΑ: Δύο σώματα έχουν ίδια μάζα (5 kg το καθένα) και βρίσκονται στη Γη. Ποια από τις παρακάτω δηλώσεις είναι σίγουρα σωστή;

- A. Έχουν ίδιο βάρος και το βάρος τους είναι 5 N
- B. Έχουν ίδιο βάρος και το βάρος τους είναι 50 N
- C. Έχουν διαφορετικό βάρος επειδή μπορεί να είναι σε διαφορετικά ύψη
- D. Το βάρος τους εξαρτάται από το σχήμα και το μέγεθός τους

ΜΑΘΗΤΗΣ: α

ΣΥΣΤΗΜΑ: Διάλεξες A. Εξήγησέ μου γιατί — τι σκέφτηκες;

ΜΑΘΗΤΗΣ: Έχουν ίδιο βάρος 5kg γιατί η μάζα και το βάρος είναι το ίδιο.

ΑΝΑΤΡΟΦΟΛΟΤΗΣΗ: ✗ Σωστή απάντηση: B

  Κύριος Γιώργος: Ακούω πάλι το ίδιο λάθος — μάζα ≠ βάρος! Μάζα (kg) = ποσότητα ύλης. Βάρος (N) = δύναμη βαρύτητας. $W = mg = 5 \times 10 = 50 \text{ N}$. Το «5» είναι η μάζα σε kg. Το «50 N» είναι το βάρος — διαφορετικοί αριθμοί με διαφορετικές μονάδες!

 V=0.0 | L=0.0 | R=0.8 | Clarity=0.8 | C= Ζωή: Αλλά στην καθημερινή γλώσσα λέμε "ζυγίζω 50 kg" — αυτό δεν είναι λάθος;

  Κύριος Γιώργος: Στα ελληνικά λέμε έτσι, αλλά στη φυσική αυτό είναι μη-ακριβές. Τεχνικά θα έπρεπε να λέμε "η μάζα μου είναι 50 kg" ή "το βάρος μου είναι 500 N".

😊 Ελισάβετ: Άρα οι διαφορετικές μάζες θα έχουν διαφορετικά βάρη;

😊 🏠 Κύριος Γιώργος: Ακριβώς! Αν το ένα σώμα είχε 10 kg και το άλλο 5 kg (ίδιος τόπος), τα βάρη θα ήταν 100 N και 50 N αντίστοιχα.

😊 🏠 Κύριος Γιώργος: Το κατάλαβες τώρα; Θες να το ξαναδούμε ή πάμε στην επόμενη ερώτηση;

ΜΑΘΗΤΗΣ: Ναι, κατάλαβα! Επόμενη →

? Ερώτηση 5/5

ΣΥΣΤΗΜΑ: Ένας αστροναύτης χρησιμοποιεί έναν ζυγό σύγκρισης (με δύο πλατφόρμες) και έναν δυναμόμετρο για να μετρήσει ένα άγνωστο αντικείμενο πρώτα στη Γη και μετά στη Σελήνη. Ποια από τις παρακάτω δηλώσεις είναι σωστή;

- A. Και τα δύο όργανα θα δείξουν ίδιες τιμές στη Γη και τη Σελήνη
- B. Ο ζυγός θα δείξει ίδια τιμή, αλλά το δυναμόμετρο θα δείξει μικρότερη τιμή στη Σελήνη
- C. Ο ζυγός θα δείξει μεγαλύτερη τιμή, αλλά το δυναμόμετρο θα δείξει ίδια τιμή
- D. Και τα δύο όργανα θα δείξουν μικρότερες τιμές στη Σελήνη

ΜΑΘΗΤΗΣ: δ

ΣΥΣΤΗΜΑ: Διάλεξες Δ (δηλαδή D). Εξήγησέ μου γιατί — τι σκέφτηκες;

ΜΑΘΗΤΗΣ: Στη Σελήνη όλα είναι πιο ελαφριά — βάρος και μάζα μειώνονται.

ΑΝΑΤΡΟΦΟΛΟΤΗΣΗ: ✗ Σωστή απάντηση: B

😊 🏠 Κύριος Γιώργος: Πάλι το ίδιο μπέρδεμα! Ζυγός σύγκρισης: συγκρίνει μάζες απευθείας → ίδιο αποτέλεσμα παντού (mg/mg το g απαλείφεται). Δυναμόμετρο: μετράει βάρος $W=mg$ → μικρότερο στη Σελήνη. Πίνακας: Ζυγός(Γη)=6kg, Ζυγός(Σελ)=6kg ✓. Δυναμόμετρο(Γη)=60N, Δυναμόμετρο(Σελ)≈10N ✗.

📊 V=0.0 | L=0.1 | R=0.7 | Clarity=0.8 | C=⌚

😊 Ζωή: Άρα ο ζυγός σύγκρισης είναι καλύτερος για ταξίδια στο διάστημα;

😊 🏠 Κύριος Γιώργος: Ακριβώς! Ο ζυγός σύγκρισης είναι αξιόπιστος παντού στο σύμπαν, γιατί δεν εξαρτάται από το g.

☹️ Ελισάβετ: Άρα αν πάμε σε έναν πλανήτη με μεγαλύτερο g (π.χ. τον Δία), το δυναμόμετρο θα δείξει πιο μεγάλη τιμή;

😊 🏠 Κύριος Γιώργος: Ναι! Στον Δία $g \approx 25 \text{ m/s}^2$ ($2.5 \times \Gamma_{\eta}$). Το ίδιο αντικείμενο θα έχει $2.5 \times$ μεγαλύτερο βάρος.

Σύνοψη ερευνητή — Σενάριο 2

Τιμές που ανέφερε το σύστημα: avg_V=0.24 | C=0.40 | avg_L=0.26 | avg_R=0.82 | avg_Clarity=0.84 | Q=0.36

Έλεγχοι συνέπειας: C1=PARTIAL(0.1) | C2=PARTIAL(0.1) | C3=FALSE(0.0) | C4=FALSE(0.0) | C5=TRUE(0.2)

Σωστές απαντήσεις: 1/5

Παράρτημα Γ : Σενάριο 3 (Μαθητής_Γ)

Σενάριο αναφοράς: επιστημονικά ορθές επιλογές και αιτιολογήσεις

Username: mathitis_C | Τάξη: Β' Γυμνασίου | Ερωτήσεις: 5

ΣΥΣΤΗΜΑ: Καλώς ήρθες στο Physics Quiz! 🏠

Βλέπω ότι θα εξεταστείς στο κεφάλαιο Μάζα και Βάρος (Γυμνάσιο).

Χρειάζομαι:

□ Ποια τάξη είσαι;

□ Πόσες ερωτήσεις θέλεις;

ΜΑΘΗΤΗΣ: Β' Γυμνασίου 5

ΣΥΣΤΗΜΑ: Τέλεια! 🎯 Είσαι Β' Γυμνασίου και θα λύσεις 5 ερωτήσεις στο κεφάλαιο
Μάζα και Βάρος.

Ας ξεκινήσουμε! 🏁

? Ερώτηση 1/5

ΣΥΣΤΗΜΑ: Ποιο από τα παρακάτω είναι σταθερό χαρακτηριστικό ενός σώματος,
ανεξάρτητα από το πού βρίσκεται;

- A. Το βάρος του
- B. Η μάζα του
- C. Η επιμήκυνση ελατηρίου που κρέμεται από αυτό
- D. Η δύναμη που ασκεί η Γη πάνω του

ΜΑΘΗΤΗΣ: B

ΣΥΣΤΗΜΑ: Διάλεξες B. Εξήγησέ μου γιατί — τι σκέφτηκες;

ΜΑΘΗΤΗΣ: Μάζα είναι η ποσότητα ύλης, που περιέχει ένα σώμα και είναι σταθερή.

ΑΝΑΤΡΟΦΟΛΟΤΗΣΗ: ✅ Σωστά! Απάντηση B είναι η σωστή!

😊 🏠 Κύριος Γιώργος: Ακριβώς! Η μάζα είναι η ποσότητα ύλης — σταθερή παντού στο
σύμπαν. Το βάρος $W=mg$ αλλάζει με το g . Στη Σελήνη ο αστροναύτης έχει ίδια μάζα αλλά
1/6 βάρος.

📊 V=1.0 | L=1.0 | R=1.0 | Clarity=1.0 | C=⌚

☺ Ζωή: Και αν πάμε στο διάστημα, μακριά από όλους τους πλανήτες — τότε τι θα συμβεί;

☺ 🏠 Κύριος Γιώργος: Η μάζα θα παραμείνει ακριβώς ίδια. Το βάρος όμως θα γίνει σχεδόν μηδέν, γιατί δεν υπάρχει βαρυτική δύναμη.

☺ 🏠 Ελισάβετ: Άρα η μάζα είναι κάτι που δεν αλλάζει ποτέ, ενώ το βάρος αλλάζει ανάλογα με τη βαρύτητα του τόπου;

☺ 🏠 Κύριος Γιώργος: Ακριβώς! Το έπιασες τέλεια. Η μάζα είναι εσωτερικό χαρακτηριστικό, το βάρος είναι δύναμη που εξαρτάται από το περιβάλλον.

☺ 🏠 Κύριος Γιώργος: Το κατάλαβες; Θες να το ξαναδούμε ή πάμε στην επόμενη ερώτηση;

ΜΑΘΗΤΗΣ: Ναι, κατάλαβα! Επόμενη →

?Ερώτηση 2/5

ΣΥΣΤΗΜΑ: Ένας μαθητής κρεμάει μια μάζα 2 kg από ένα ελατήριο και μετράει την επιμήκυνσή του. Αν την ίδια μάζα την κρεμούσε στη Σελήνη ($g_{\text{σελ}}=g/6$) από το ίδιο ελατήριο, τι θα συμβεί στην επιμήκυνση;

- A. Θα παραμείνει ίδια
- B. Θα αυξηθεί
- C. Θα μειωθεί
- D. Θα μηδενιστεί

ΜΑΘΗΤΗΣ: γ

ΣΥΣΤΗΜΑ: Διάλεξες Γ (C). Εξήγησέ μου γιατί — τι σκέφτηκες;

ΜΑΘΗΤΗΣ: Γιατί η επιμήκυνση του ελατηρίου εξαρτάται από το βάρος του σώματος και όχι την μάζα, οπότε επειδή το βάρος στη Σελήνη είναι μικρότερο, άρα και η επιμήκυνση θα είναι μικρότερη.

ΑΝΑΤΡΟΦΟΔΟΤΗΣΗ: ☒ Σωστά! Απάντηση C είναι η σωστή!

☺ 🏠 Κύριος Γιώργος: Εξαιρετική σκέψη! $x = mg/k$. Σελήνη: g μικρότερο $\rightarrow W$ μικρότερο $\rightarrow x$ μικρότερο. Γη: $W=2 \times 10=20\text{N}$. Σελήνη: $W=2 \times 1.6=3.2\text{N}$. Η κλίση της γραφικής $m-x$ αλλάζει αλλά η γραμμικότητα παραμένει!

📊 V=1.0 | L=1.0 | R=1.0 | Clarity=1.0 | C=⌚

☺ Ζωή: Και αν βάζαμε ένα ακόμα βαρύτερο ελατήριο στη Γη, θα έχουμε ακόμα μεγαλύτερη επιμήκυνση για ίδια μάζα;

☺ 🏠 Κύριος Γιώργος: Όχι! Βαρύτερο (πιο σκληρό) ελατήριο έχει μεγαλύτερο $k \rightarrow$ μικρότερη επιμήκυνση για ίδια μάζα.

☺ Ελισάβετ: Άρα αν αλλάξουμε ελατήριο ή αλλάξουμε πλανήτη, και τα δύο αλλάζουν την επιμήκυνση αλλά για διαφορετικούς λόγους;

☺ 🏠 Κύριος Γιώργος: Ακριβώς! Αλλαγή ελατηρίου = αλλαγή k . Αλλαγή πλανήτη = αλλαγή g . Και οι δύο αλλάζουν την επιμήκυνση.

☺ 🏠 Κύριος Γιώργος: Πολύ καλά! Συνεχίζουμε! 📖

?Ερώτηση 3/5

ΣΥΣΤΗΜΑ: Δύο σώματα A και B έχουν μάζες $m_A=2m_B$. Και τα δύο βρίσκονται στη Γη ($g=10 \text{ m/s}^2$). Ποιος είναι ο λόγος των βαρών τους W_A/W_B ;

- A. $W_A/W_B = 4$
- B. $W_A/W_B = 1/2$
- C. $W_A/W_B = 2$
- D. $W_A/W_B = 1/4$

ΜΑΘΗΤΗΣ: γ

ΣΥΣΤΗΜΑ: Διάλεξες Γ (C). Εξήγησέ μου γιατί — τι σκέφτηκες;

ΜΑΘΗΤΗΣ: Επειδή η επιτάχυνση βαρύτητας είναι ίδια για όλα τα σώματα στον ίδιο τόπο, άρα ο λόγος των βαρών θα είναι ίδιος με τον λόγο των μαζών.

ΑΝΑΤΡΟΦΟΛΟΤΗΣΗ: ✓ Σωστά! Απάντηση C είναι η σωστή!

☺ 🏠 Κύριος Γιώργος: Εξαιρετικό! $W_A/W_B = m_A \cdot g / m_B \cdot g = m_A/m_B = 2$. Το g απαλείφεται! Στη Σελήνη ο λόγος παραμένει 2 — τα απόλυτα βάρη αλλάζουν αλλά ο λόγος ίδιος.

📊 $V=1.0 \mid L=1.0 \mid R=1.0 \mid \text{Clarity}=1.0 \mid C=\text{🕒}$

☺ Ζωή: Και αν τα δύο σώματα ήταν στη Σελήνη αντί για τη Γη, θα ήταν ο λόγος διαφορετικός;

  Κύριος Γιώργος: Όχι! Ο λόγος θα παρέμενε ακριβώς 2. Το $g_{\text{σελ}}$ απαλείφεται επίσης στον λόγο.

 Ελισάβετ: Άρα ο λόγος δύο μαζών είναι πάντα ίδιος με τον λόγο των βαρών τους, ό,τι και να αλλάξει το g ;

  Κύριος Γιώργος: Ναι, εφόσον είναι στον ίδιο τόπο! Αν ήταν σε διαφορετικούς τόπους (Α στη Γη, Β στη Σελήνη), τότε ο λόγος βαρών θα ήταν διαφορετικός από τον λόγο μαζών.

  Κύριος Γιώργος: Υπέροχα! Συνεχίζουμε! 

?Ερώτηση 4/5

ΣΥΣΤΗΜΑ: Ένας αστροναύτης έχει μάζα 80 kg. Στη Γη το βάρος του είναι 800 N ($g=10 \text{ m/s}^2$). Όταν φτάσει στη Σελήνη ($g_{\text{σελ}}=1.6 \text{ m/s}^2$), ποιες δηλώσεις είναι σωστές;

I. Η μάζα παραμένει 80 kg

II. Το βάρος στη Σελήνη είναι 128 N

III. Ο λόγος $W_{\text{γη}}/W_{\text{σελ}} = 6.25$

A. Μόνο I B. Μόνο II C. I+II D. Όλες

ΜΑΘΗΤΗΣ: δ

ΣΥΣΤΗΜΑ: Διάλεξες Δ (D). Εξήγησέ μου γιατί — τι σκέφτηκες;

ΜΑΘΗΤΗΣ: Η μάζα παραμένει ίδια και το βάρος στη Σελήνη είναι περίπου το 1/6 της Γης.

ΑΝΑΤΡΟΦΟΔΟΤΗΣΗ:  Σωστά! Απάντηση D είναι η σωστή!

  Κύριος Γιώργος: Εξαιρετικά! I: μάζα=80kg παντού ✓. II: $W=80 \times 1.6=128\text{N}$ ✓. III: $800/128=6.25$ ✓. Προσέχουμε: «1/6 βάρος» → λόγος=6.25 (όχι ακριβώς 6, γιατί $g_{\text{σελ}}=1.6$, όχι $g/6=1.667$).

 V=1.0 | L=0.9 | R=1.0 | Clarity=0.9 | C=

 Ζωή: Άρα ο αστροναύτης θα ήταν ελαφρύτερος στη Σελήνη — θα μπορούσε να πηδάει πολύ ψηλά;

  Κύριος Γιώργος: Ακριβώς! Μικρότερο βάρος → πιο εύκολη κίνηση. Αλλά η αδράνεια (αντίσταση στη μεταβολή κίνησης) παραμένει ίδια, γιατί εξαρτάται από τη μάζα!

 Ελισάβετ: Άρα χρειάζεται ίδια δύναμη για να αλλάξει την κίνησή του, ακόμα κι αν το βάρος είναι μικρότερο;

  Κύριος Γιώργος: Τέλεια παρατήρηση! Η αδράνεια εξαρτάται από τη μάζα — εντελώς διαφορετικό από το βάρος!

  Κύριος Γιώργος: Πάμε στην τελευταία ερώτηση! 

?Ερώτηση 5/5

ΣΥΣΤΗΜΑ: Ένας μαθητής χρησιμοποιεί ζυγό σύγκρισης και δυναμόμετρο για να μετρήσει ένα άγνωστο σώμα. Ποιες δηλώσεις είναι σωστές;

- I. Ο ζυγός σύγκρισης θα δείξει το ίδιο αποτέλεσμα στη Γη και στη Σελήνη
II. Το δυναμόμετρο θα δείξει διαφορετικό αποτέλεσμα στη Γη και στη Σελήνη
III. Ο ζυγός μετράει μάζα, ενώ το δυναμόμετρο μετράει βάρος
A. Μόνο I B. Μόνο II C. I+II D. Όλες

ΜΑΘΗΤΗΣ: δ

ΣΥΣΤΗΜΑ: Διάλεξες Δ (D). Εξήγησέ μου γιατί — τι σκέφτηκες;

ΜΑΘΗΤΗΣ: Ο ζυγός σύγκρισης μετρά τη μάζα, που είναι παντού ίδια. Το δυναμόμετρο μετρά το βάρος που αλλάζει.

ΑΝΑΤΡΟΦΟΔΟΤΗΣΗ:  Σωστά! Απάντηση D είναι η σωστή!

  Κύριος Γιώργος: Εξαιρετική σύνοψη! I: Ζυγός = μάζες (το g απαλείφεται) → ίδιο παντού ✓. II: Δυναμόμετρο = βάρος $W=mg$ → μικρότερο στη Σελήνη ✓. III: Θεμελιώδης διάκριση: ζυγός=kg, δυναμόμετρο=N ✓.

 V=1.0 | L=1.0 | R=1.0 | Clarity=1.0 | C=

 Ζωή: Άρα αν πήγαμε στη Σελήνη με ηλεκτρονικό ζυγό, θα δείχνε λάθος;

  Κύριος Γιώργος: Ναι! Ο ηλεκτρονικός ζυγός μετράει βάρος εσωτερικά και διαιρεί με g_γη. Στη Σελήνη, υπολογίζει λάθος μάζα!

 Ελισάβετ: Άρα ο ζυγός σύγκρισης είναι πιο αξιόπιστος αν θέλουμε μάζα σε διαφορετικούς πλανήτες;

  Κύριος Γιώργος: Εντελώς σωστό! Ο ζυγός σύγκρισης είναι καθολικός — λειτουργεί παντού.

  Κύριος Γιώργος: Συγχαρητήρια! Τελείωσες όλες τις ερωτήσεις! 

Σύνοψη ερευνητή — Σενάριο 3

**Τιμές που ανέφερε το σύστημα: avg_V=1.0 | C=1.0 | avg_L=0.98 | avg_R=1.0 |
avg_Clarify=0.98 | Q=1.0**

Έλεγχοι συνέπειας: C1=TRUE(0.2) | C2=TRUE(0.2) | C3=TRUE(0.2) | C4=TRUE(0.2) |
C5=TRUE(0.2)

Σωστές απαντήσεις: 5/5

Παράρτημα Δ : Δομές Q_RECORD και QUIZ_SUMMARY

Στο παράρτημα αυτό παρατίθενται οι δομές των εγγραφών Q_RECORD και QUIZ_SUMMARY, όπως ορίζονται στο system prompt του Ph-D A.I. (PhysicsDialog AI). Πρόκειται για πρότυπα συμπλήρωσης από το μοντέλο και όχι για εξαγωγή των πραγματικών εγγραφών των τριών σεναρίων. Τα Q_RECORD αφαιρούνται από το εμφανιζόμενο κείμενο στη διεπαφή μέσω JavaScript, ενώ το QUIZ_SUMMARY αφαιρείται από την απόκριση από το PHP πριν αυτή εμφανιστεί στον χρήστη. Η εξαγωγή, η αποθήκευση και οι περιορισμοί των πεδίων τους περιγράφονται στις Ενότητες 3.8, 3.10 και 4.4. Η Ενότητα Δ.3 παρουσιάζει τη μορφή με την οποία το PHP αποθηκεύει τη σύνοψη στον διακομιστή. Οι τιμές σε αγκύλες αποτελούν θέσεις συμπλήρωσης και δεν συνιστούν έγκυρο JSON πριν αντικατασταθούν με τις αντίστοιχες τιμές.

Δ.1 Q_RECORD — εγγραφή ανά ερώτηση

```
%%Q_RECORD_START%%
{
  "q_num": [N],
  "question_text": "[ακριβές κείμενο ερώτησης]",
  "options": {"A": "[...]", "B": "[...]", "C": "[...]", "D": "[...]"},
  "correct_choice": "[X]",
  "user_choice": "[X]",
  "is_correct": [true/false],
  "user_explanation": "[αυτολεξεί η εξήγηση του μαθητή – μην  
παραφράξεις]",
  "summary": "[Περίληψη: τι απάντησε ο μαθητής, η βασική λογική της  
αιτιολόγησής του και ανάλυση σωστού/λάθους]",
  "scores": {
    "V": [0.x],
    "L": [0.x],
    "R": [0.x],
    "Clarity": [0.x],
    "C": "pending"
  },
  "score_analysis": {
    "V": {
      "value": [0.x],
      "justification": "[Γιατί πήρε αυτή την τιμή – συγκεκριμένα από τα  
V1-V5 checks]"
    },
    "L": {
      "value": [0.x],
      "justification": "[Γιατί πήρε αυτή την τιμή – συγκεκριμένα από τα  
L1-L5 checks]"
    },
    "R": {
      "value": [0.x],
      "justification": "[Γιατί – π.χ. 'χρησιμοποίησε τύπο εκτός βαθμίδας'  
ή 'κατάλληλο επίπεδο']"
```

```

    },
    "Clarity": {
        "value": [0.x],
        "justification": "[Γιατί - π.χ. 'σαφής δομή αλλά ασαφής ορολογία
στη 2η πρόταση']"
    },
    "C": {
        "value": "pending",
        "justification": "Υπολογίζεται στο τέλος βάσει ολόκληρου κουίζ"
    }
},
"q_score": "pending",
"re_explain": "pending"
}
%%Q_RECORD_END%%

```

Δ.2 QUIZ_SUMMARY — σύνοψη συνεδρίας

```

%%QUIZ_SUMMARY_START%%
{
    "topic": "[topic]",
    "grade": "[grade]",
    "num_questions": [N],
    "correct_answers": [N_correct],
    "avg_V": [0.x],
    "avg_L": [0.x],
    "avg_R": [0.x],
    "avg_Clarify": [0.x],
    "C_score": [0.x],
    "C_checks": {
        "C1": [0.x],
        "C2": [0.x],
        "C3": [0.x],
        "C4": [0.x],
        "C5": [0.x]
    },
    "avg_Q": [0.x],
    "strong_points": ["...", "..."],
    "weak_points": ["...", "..."],
    "re_explain_questions": [αριθμοί ή κενό array]
}
%%QUIZ_SUMMARY_END%%

```

Δ.3 Μόνιμη εγγραφή στον διακομιστή

Όταν εντοπιστεί αναγνώσιμο QUIZ_SUMMARY, το PHP προσθέτει μία νέα εγγραφή στον πίνακα JSON του αρχείου logs/physics_quiz/history/{username}.json. Το session_id και το timestamp δημιουργούνται από το PHP και όχι από τη διεπαφή. Το πεδίο questions λαμβάνει τιμή από το αντίστοιχο πεδίο του QUIZ_SUMMARY· επειδή το σχήμα του prompt (Δ.2) δεν περιλαμβάνει τέτοιο πεδίο, μένει κενό (Ενότητα 3.8.2).

```

[
  {
    "session_id": "[username]_[EEEE-MM-HH_ωω-λλ-δδ]_[6 χαρακτήρες]",

```

```
"timestamp": "[ημερομηνία και ώρα σε μορφή ISO 8601]",  
"username": "[username]",  
"summary": { [όλα τα πεδία του QUIZ_SUMMARY της Δ.2] },  
"questions": []  
},  
{ [επόμενη εγγραφή] }  
]
```

Υπεύθυνη Δήλωση Συγγραφέα:

Δηλώνω ρητά ότι, σύμφωνα με το άρθρο 8 του Ν.1599/1986, η παρούσα εργασία αποτελεί αποκλειστικά προϊόν προσωπικής μου εργασίας, δεν προσβάλλει κάθε μορφής δικαιώματα διανοητικής ιδιοκτησίας, προσωπικότητας και προσωπικών δεδομένων τρίτων, δεν περιέχει έργα/εισφορές τρίτων για τα οποία απαιτείται άδεια των δημιουργών/δικαιούχων και δεν είναι προϊόν μερικής ή ολικής αντιγραφής, οι πηγές δε που χρησιμοποιήθηκαν περιορίζονται στις βιβλιογραφικές αναφορές και μόνον και πληρούν τους κανόνες της επιστημονικής παράθεσης.