

Η Συνεισφορά της Μηχανικής Προτροπών με Μεγάλα Γλωσσικά Μοντέλα στην Ανάπτυξη Δεξιοτήτων Επιστημονικού Συλλογισμού στο Μάθημα της Φυσικής

Αντώνιος Ματσίγκος¹ και Γεώργιος Κρητικός²

¹Υποψήφιος Διδάκτορας, ²Εργαστηριακό Διδακτικό Προσωπικό,

^{1,2}Τμήμα Επιστημών της Προσχολικής Αγωγής και του Εκπαιδευτικού Σχεδιασμού,
Πανεπιστήμιο Αιγαίου

¹matsigkosa@gmail.com, ²gkritikos@aegean.gr

Περίληψη

Η παρούσα έρευνα εξετάζει τον ρόλο της μηχανικής προτροπών (prompt engineering) με τα μεγάλα γλωσσικά μοντέλα (Large Language Models-LLMs), όπως το ChatGPT-4, ως μαθησιακό εργαλείο για την ανάπτυξη δεξιοτήτων επιστημονικού συλλογισμού και την ενίσχυση της ικανότητας επίλυσης προβλημάτων Φυσικής σε μαθητές δευτεροβάθμιας εκπαίδευσης. Βασικός στόχος της έρευνας είναι να ενσωματώσει τα μοντέλα τεχνητής νοημοσύνης στις παραδοσιακές μεθόδους διδασκαλίας, ως «υποστηρικτές» της μάθησης σε ένα μαθητοκεντρικό περιβάλλον. Αξιοποιώντας στρατηγικές της μηχανικής προτροπών, όπως η αλυσίδα σκέψης (Chain of Thought-CoT), επιδιώκεται η ενίσχυση της κριτικής σκέψης των μαθητών και η κατανόηση θεμελιωδών αρχών της Φυσικής.

Λέξεις κλειδιά: δεξιότητες επιστημονικού συλλογισμού, επίλυση προβλήματος, μεγάλα γλωσσικά μοντέλα, μηχανική προτροπών

The Contribution of Prompt Engineering with Large Language Models to the Development of Scientific Reasoning Skills in Physics Courses

Antonios Matsigkos¹ and Georgios Kritikos²

¹PhD Student, ²Laboratory Teaching Staff,

^{1,2}Department of Preschool Education Sciences and Educational Design,
University of the Aegean

¹matsigkosa@gmail.com, ²gkritikos@aegean.gr

Abstract

This research examines the role of prompt engineering with Large Language Models (LLMs), such as ChatGPT-4, as a learning tool for developing scientific reasoning skills and enhancing the ability to solve Physics problems for secondary school students. The main goal of the research is to integrate artificial intelligence models into traditional teaching methods, as "supporters" of learning in a student-centered environment. Utilizing prompt engineering strategies such as the chain of thought (CoT), the aim is to strengthen students' critical thinking and understanding of fundamental principles of Physics.

Keywords: Large Language Models (LLMs), problem solving, prompt engineering, scientific reasoning skills

Εισαγωγή

Η άνοδος της τεχνητής νοημοσύνης και οι εφαρμογές της αναπόφευκτα εισβάλλουν σε διάφορους τομείς της ζωής του ανθρώπου, συμπεριλαμβανομένης της εκπαίδευσης. Στην εκπαίδευση των φυσικών επιστημών, η ανάπτυξη δεξιοτήτων επιστημονικού συλλογισμού θεωρείται κρίσιμη για την κατανόηση και την επίλυση πολύπλοκων προβλημάτων (Lawson, 2000). Η συγκεκριμένη έρευνα στοχεύει στην ανάπτυξη μιας στρατηγικής επίλυσης προβλημάτων Φυσικής, η οποία προωθεί την ανάπτυξη βασικών δεξιοτήτων επιστημονικού συλλογισμού. Ο σχεδιασμός αλυσίδας σκέψης (CoT) βασισμένης σε δεξιότητες επιστημονικού συλλογισμού από τον εκπαιδευτικό και η υιοθέτησή της από το μοντέλο (LLM) μέσα από την μηχανική προτροπών (prompt engineering), διαμορφώνει ένα υποστηρικτικό πλαίσιο επίλυσης προβλημάτων Φυσικής εστιάζοντας στην αποτελεσματική και παραγωγική χρήση της τεχνητής νοημοσύνης.

Θεωρητικό Πλαίσιο

Στη μάθηση με δυνατότητα τεχνητής νοημοσύνης, η εμφάνιση μεγάλων γλωσσικών μοντέλων όπως το ChatGPT και το davinci-003 έχει δείξει πιθανά οφέλη στην εκπαίδευση στη Φυσική (MacIsaac, 2023 · Yeadon et al., 2022). Αν και αρχικά έρευνες είχαν διαπιστώσει πως τα LLMs, ενώ είχαν καλή απόδοση σε εργασίες επεξεργασίας φυσικής γλώσσας και υστερούσαν σημαντικά σε παραδείγματα εργασιών που απαιτούσαν μαθηματικές επεξεργασίες, μετά την κυκλοφορία των γνωστών LLMs-powered chatbot ChatGPT, όπως το GPT-4, φαίνεται πως πλέον μπορούν να επιλύουν σύνθετες εργασίες σε επίπεδα απόδοσης παρόμοια με αυτά του ανθρώπου (Bubeck et al., 2023).

Η Μηχανική Προτροπών (Prompt Engineering) είναι ένα πεδίο-τομέας που περιλαμβάνει πολλαπλές τεχνικές και στρατηγικές για τη βελτίωση της απόδοσης των γλωσσικών μοντέλων (LLMs – Large Language Models). Προτροπή ορίζεται η οδηγία φυσικής γλώσσας που δίνεται σε ένα LLM για να διευκολυνθεί η δημιουργία μίας εξόδου και αποτελεί μία μορφή «άμεσου προγραμματισμού» (White et al., 2024). Στην παρούσα εργασία ερευνάται ο αντίκτυπος της μηχανικής προτροπών, μέσα από τη χρήση τεχνικών όπως είναι η προτροπή μάθησης λίγων βολών (Few-Shot Prompting) και η προτροπή αλυσίδας σκέψης (Chain of Thought-CoT) (Sahoo et al., 2024 · Wei et al., 2022) με σκοπό τη βελτίωση της απόδοσης του μοντέλου στον τομέα της Φυσικής και τη διαμόρφωση μαθησιακού υποστηρικτικού περιβάλλοντος.

Τα τελευταία χρόνια η ενσωμάτωση Μεγάλων Γλωσσικών Μοντέλων (LLMs) στην εκπαιδευτική διαδικασία των Φυσικών Επιστημών αποκτά ολοένα και μεγαλύτερο ερευνητικό ενδιαφέρον, καθώς προκύπτουν ενδείξεις για τη θετική τους συνεισφορά στην κατανόηση εννοιών και στην ανάπτυξη δεξιοτήτων επιστημονικού συλλογισμού. Η έρευνα των Jiang και Jiang (2024) αναδεικνύει τη δυνατότητα αξιοποίησης Μεγάλων Γλωσσικών Μοντέλων (LLMs) ως εργαλείων εξατομικευμένης υποστήριξης στη διδασκαλία της Φυσικής. Συγκεκριμένα, μέσω της ανάπτυξης ενός συστήματος καθοδήγησης που βασίζεται σε LLM, οι ερευνητές διαπίστωσαν βελτίωση στη βαθύτερη κατανόηση των φυσικών εννοιών και ενίσχυση των δεξιοτήτων επιστημονικού συλλογισμού των μαθητών. Η αλληλεπίδραση με το σύστημα δεν περιορίστηκε στη σωστή απάντηση, αλλά επικεντρώθηκε στη σταδιακή οικοδόμηση εννοιολογικής κατανόησης και στην ανάπτυξη ορθών αιτιακών συλλογισμών.

Μεθοδολογία έρευνας

Η συγκεκριμένη έρευνα περιλαμβάνει δύο φάσεις. Η πρώτη φάση αφορά στη βελτιστοποίηση της απόδοσης του μεγάλου γλωσσικού μοντέλου (LLM), και η δεύτερη στην αλληλεπίδραση μοντέλου-μαθητή για την επίλυση προβλήματος Φυσικής. Η επίλυση προβλήματος στην παρούσα πιλοτική εργασία, σχετίζεται με τη μελέτη της οριζόντιας βολής σώματος. Η μελέτη του συγκεκριμένου φαινομένου προσφέρεται τόσο για δυναμική-κινηματική προσέγγιση αξιοποιώντας τους νόμους του Νεύτωνα και τις εξισώσεις κίνησης, όσο και για ενεργειακή

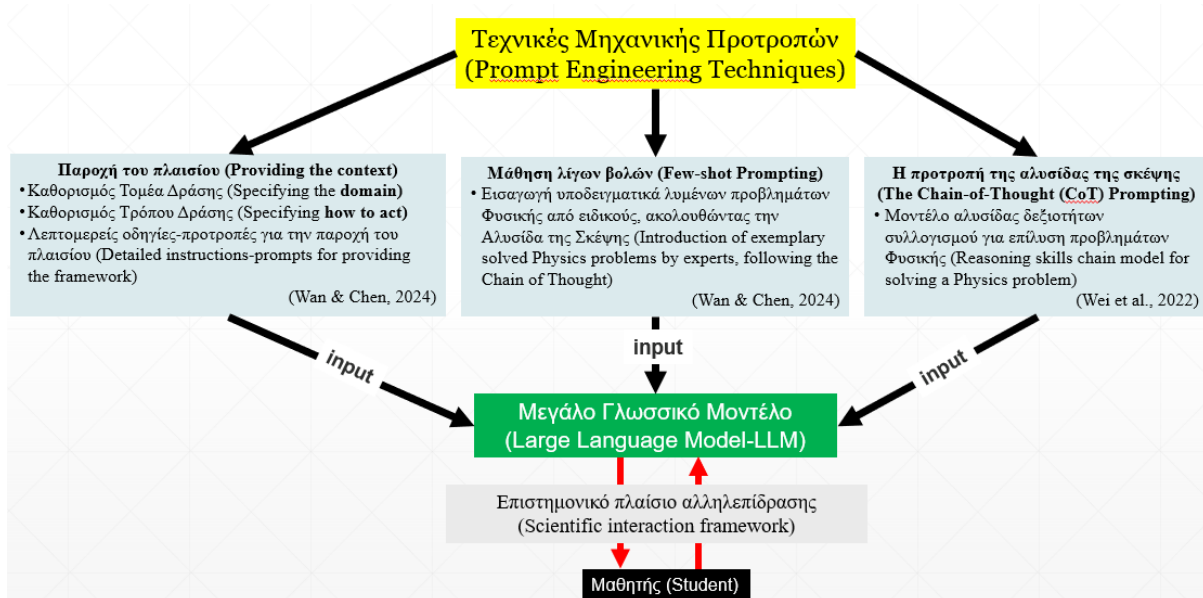
εφαρμόζοντας την Αρχή Διατήρησης της Ενέργειας. Οι δύο αυτές προσεγγίσεις είναι θεμελιώδεις για την επίλυση προβλήματος στην κλασική Φυσική.

Το δείγμα της πιλοτικής έρευνας ήταν 3 μαθητές (2 αγόρια και ένα κορίτσι) της Β' τάξης Γενικού Λυκείου από την Ομάδα Θετικών Σπουδών. Η επιλογή ενός μικρού πιλοτικού δείγματος είναι συνεπής με τη φύση της πιλοτικής έρευνας, η οποία αποσκοπεί στην αξιολόγηση της σκοπιμότητας της παρέμβασης και στην αναγνώριση τυχόν προκλήσεων πριν από μια μελέτη μεγαλύτερης κλίμακας. Οι μαθητές επιλέχθηκαν με βάση την εθελοντική τους συμμετοχή. Πριν την έναρξη της έρευνας, εξασφαλίστηκε η ενημερωμένη συγκατάθεση τόσο των μαθητών όσο και των κηδεμόνων τους, διασφαλίζοντας παράλληλα την πλήρη ανωνυμία και εμπιστευτικότητα των δεδομένων που συλλέχθηκαν.

Κατά την πρώτη φάση διαμορφώθηκαν κατάλληλες προτροπές με σκοπό την «εκπαίδευση» του μοντέλου. Αρχικά, εισήχθη μία λίστα προτροπών-οδηγιών στο μοντέλο, η οποία είχε ως στόχο τη διαμόρφωση επιστημονικού πλαισίου αλληλεπίδρασης με τον χρήστη-μαθητή (Εικόνα 1). Το πρώτο στοιχείο της προτροπής αφορά στην παροχή του πλαισίου με στρατηγικές της μηχανικής προτροπών, όπως ο καθορισμός του τομέα δράσης και ο καθορισμός του τρόπου δράσης. Το δεύτερο στοιχείο που εισήχθη στο μοντέλο είναι η δομή της αλυσίδας σκέψης (8 γνωστικών σταδίων), βασισμένη σε θεμελιώδεις δεξιότητες επιστημονικού συλλογισμού, όπως φαίνεται στον Πίνακα 1. Τέλος, αξιοποιήθηκε η τεχνική της μάθησης λίγων βολών (Few-shot Prompting) εισάγοντας στο μοντέλο υποδειγματικά λυμένα προβλήματα Φυσικής, με σκοπό να υιοθετήσει την τεχνοτροπία της αλυσίδας σκέψης πάνω στην επίλυση προβλήματος.

Στη συνέχεια, ελέγχθηκε η ποιότητα των ανατροφοδοτήσεων του μοντέλου προς τον χρήστη από ομάδα ειδικών (καθηγητές Φυσικής) και έγιναν κατάλληλες τροποποιήσεις στην αρχιτεκτονική των προτροπών ώστε να εξασφαλιστεί ικανοποιητική απόδοση σε θέματα καθοδήγησης επίλυσης προβλήματος.

Σχήμα 1. Σχήμα “Εκπαίδευσης” του μοντέλου με χρήση τεχνικών του Prompt Engineering



Η κεντρική φάση της έρευνας αφορούσε στην εκπαιδευτική παρέμβαση, κατά την οποία οι μαθητές αλληλεπίδρασαν με το Μεγάλο Γλωσσικό Μοντέλο (LLM) ChatGPT-4. Να σημειωθεί πως η παρέμβαση πραγματοποιήθηκε μετά την ολοκλήρωση της διδακτέας ύλης με την παραδοσιακή μέθοδο (διάλεξη καθηγητή, συζήτηση στην τάξη, επίδειξη εφαρμογών, επίλυση σχετικών προβλημάτων) και κατά την περίοδο των επαναλήψεων. Αυτή η αλληλεπίδραση πραγματοποιήθηκε εξατομικευμένα, με το μοντέλο να λειτουργεί ως έμπειρος εκπαιδευτικός

Φυσικής, εστιάζοντας στην επίλυση προβλημάτων του φαινομένου της οριζόντιας βολής. Το ChatGPT-4 καθοδηγούσε τους μαθητές μέσω της δομημένης αλυσίδας σκέψης (Chain of Thought - CoT) 8 σταδίων του Πίνακα 1, η οποία είχε ενσωματωθεί στην προτροπή (prompt) του LLM, με τεχνικές Prompt Engineering. Περιλαμβάνει δεξιότητες όπως η αναγνώριση των φαινομένων που πραγματεύεται το προς επίλυση πρόβλημα (Chi et al., 1981· Polya, 2014), η αναγνώριση αρχικών και τελικών συνθηκών του συστήματος (Chi et al., 1981), η διατύπωση υποθέσεων και αναγνώριση αιτιακών σχέσεων (Lawson, 2000), η αναγνώριση και ο έλεγχος των μεταβλητών (Chi et al., 1981· Lawson, 1978), ο συσχετιστικός συλλογισμός, ο παραγωγικός συλλογισμός, η εξαγωγή αποτελεσμάτων και η αξιολόγηση (Zimmerman, 2007).

Η αλληλεπίδραση δομήθηκε μέσω ερωταποκρίσεων, με το LLM να θέτει μία ερώτηση τη φορά. Εφαρμόστηκε προσαρμοστική καθοδήγηση (scaffolding), όπου η βοήθεια μειωνόταν σταδιακά ανάλογα με την ορθότητα των απαντήσεων του μαθητή (π.χ. μετά από δύο διαδοχικές σωστές απαντήσεις). Επιπλέον, παρέχόταν προσαρμοστική ανατροφοδότηση (adaptive feedback): σε περίπτωση λανθασμένης απάντησης, το μοντέλο δεν έδινε τη σωστή λύση, αλλά προέτρεπε τον μαθητή να επανεξετάσει τη σκέψη του, υιοθετώντας πρακτικές Σωκρατικού διαλόγου. Η δυσκολία και η μορφή των ερωτήσεων προσαρμόζονταν επίσης δυναμικά στην απόδοση του μαθητή. Κάθε μαθητής συμμετείχε σε τρεις (3) διαφορετικές αλληλεπιδράσεις, με μέση χρονική διάρκεια 30 λεπτά ανά συνεδρία. Καθ' όλη τη διάρκεια των αλληλεπιδράσεων, υπήρχε επίβλεψη από εκπαιδευτικό, ο οποίος διασφάλιζε την ομαλή διεξαγωγή και παρείχε τεχνική υποστήριξη, χωρίς να παρεμβαίνει στο περιεχόμενο της καθοδήγησης του LLM.

Τα ερευνητικά ερωτήματα που τέθηκαν για τη διερεύνηση αυτής της πιλοτικής μελέτης είναι τα εξής:

1. Πώς η καθοδηγούμενη αλληλεπίδραση μεταξύ μαθητή και Μεγάλου Γλωσσικού Μοντέλου (LLM), με χρήση τεχνικών prompt engineering, επηρεάζει την ανάπτυξη δεξιοτήτων επίλυσης προβλήματος και επιστημονικού συλλογισμού στη Φυσική;
2. Πώς αντιλαμβάνονται οι μαθητές την καθοδηγούμενη αλληλεπίδραση με Μεγάλο Γλωσσικό Μοντέλο (LLM) στην επίλυση προβλημάτων οριζόντιας βολής μέσω μιας δομημένης αλυσίδας σκέψης, και κατά πόσο θεωρούν ότι αυτή υποστηρίζει την κατανόηση των εννοιών και την ανάπτυξη δεξιοτήτων επιστημονικού συλλογισμού;

Πίνακας 1. Αλυσίδα σκέψης (Chain of Thought-CoT) για την επίλυση προβλήματος Φυσικής

1. Αναγνώριση φαινομένων του προβλήματος (Phenomenon Recognition)
2. Οπτικοποίηση και σχεδιασμός (Visualization and design)
3. Εντοπισμός των κεντρικών ερωτημάτων του προβλήματος (Identifying the central questions of the problem)
4. Διάκριση τύπου φαινομένου (Στατικά-Δυναμικά φαινόμενα) (Phenomenon Classification)
5. Εντοπισμός μεταβλητών του προβλήματος - Συσχετιστικός συλλογισμός (Control of Variables - Associative Reasoning)
6. Διατύπωση υποθέσεων - Αιτιακός συλλογισμός (Hypothesis Formulation - Causal Reasoning) - Ανάλυση δυναμικής και ενεργειακής εξέλιξης του φαινομένου (Phenomenon Dynamic & Energy Analysis)
7. Εφαρμογή των φυσικών αρχών που διέπουν το φαινόμενο (Application of the physical principles governing the phenomenon)
8. Έλεγχος και αξιολόγηση αποτελεσμάτων (Validation and Evaluation of Results)

Σημείωση: Στον πίνακα 1, αναγράφονται σε τίτλους βασικά θεμελιώδη βήματα επιστημονικού συλλογισμού για την επίλυση προβλήματος Φυσικής, τα οποία υιοθετούνται στις προτροπές προς το μοντέλο (LLM).

Για τη διερεύνηση των ερευνητικών ερωτημάτων και την αξιολόγηση της επίδρασης της εκπαιδευτικής παρέμβασης, χρησιμοποιήθηκαν δύο ερωτηματολόγια συλλογής δεδομένων, μέσω της πλατφόρμας Google Forms. Πρωτίστως, αναπτύχθηκε ένα ερωτηματολόγιο προ-ελέγχου και μετά-ελέγχου 19 ερωτήσεων πολλαπλής επιλογής (pre-test/post-test, βλ. Παράρτημα), το οποίο σχεδιάστηκε με σκοπό την ποσοτική αξιολόγηση της ανάπτυξης

συγκεκριμένων δεξιοτήτων επιστημονικού συλλογισμού στους συμμετέχοντες μαθητές. Το ερωτηματολόγιο συμπληρώθηκε από τους μαθητές πριν και μετά την παρέμβαση.

Οι δεξιότητες που διερευνήθηκαν στην παρούσα πιλοτική έρευνα, σύμφωνα με τα ερευνητικά ερωτήματα, περιλάμβαναν:

1. Αναγνώριση Φαινομένου (Phenomenon Recognition) (ερωτήσεις 1-5): Αφορά στην ικανότητα των μαθητών να αναγνωρίζουν και να κατανοούν το φυσικό φαινόμενο που περιγράφεται στο πρόβλημα.

2. Έλεγχος Μεταβλητών - Συσχετιστικός Συλλογισμός (Control of Variables - Associative Reasoning) (ερωτήσεις 6-14): Εστιάζει στην ικανότητα αναγνώρισης και συσχέτισης των φυσικών ποσοτήτων (φυσικών μεγεθών) που εμπλέκονται στο φαινόμενο. Εδώ δημιουργήθηκαν 4 διττές ερωτήσεις (7-14). Στην πρώτη ερώτηση κάθε ζεύγους ελέγχεται η άποψη των μαθητών για την εξάρτηση ορισμένων Φυσικών μεγεθών από μεταβλητές (συσχέτιση μεταβλητών), και ακολουθεί μια δεύτερη ερώτηση η οποία αποτελεί εφαρμογή ελέγχου της κατανόησης της σχέσης των μεταβλητών.

3. Διατύπωση Υποθέσεων - Αιτιακός Συλλογισμός (Hypothesis Formulation - Causal Reasoning) (ερωτήσεις 15-19): Αξιολογεί την ικανότητα του μαθητή να αναλύει τη δυναμική και ενεργειακή εξέλιξη του φαινομένου, να διατυπώνει υποθέσεις και να ανιχνεύει τις σχέσεις αιτίας αποτελέσματος (π.χ. η βαρυτική δύναμη είναι η αιτία για τη μείωση της βαρυτικής δυναμικής ενέργειας του σώματος στο φαινόμενο της οριζόντιας βολής).

Δευτερευόντως, χρησιμοποιήθηκε ένα ερωτηματολόγιο αποτίμησης της μαθησιακής εμπειρίας, δομημένο σε πενταβάθμια κλίμακα Likert. Σκοπός του ήταν η καταγραφή των υποκειμενικών αντιλήψεων των μαθητών σχετικά με την καθοδηγητική αλληλεπίδρασή τους με το Μεγάλο Γλωσσικό Μοντέλο (LLM), καθώς και του βαθμού στον οποίο θεώρησαν ότι αυτή η παρέμβαση συνέβαλε στη μαθησιακή τους υποστήριξη.

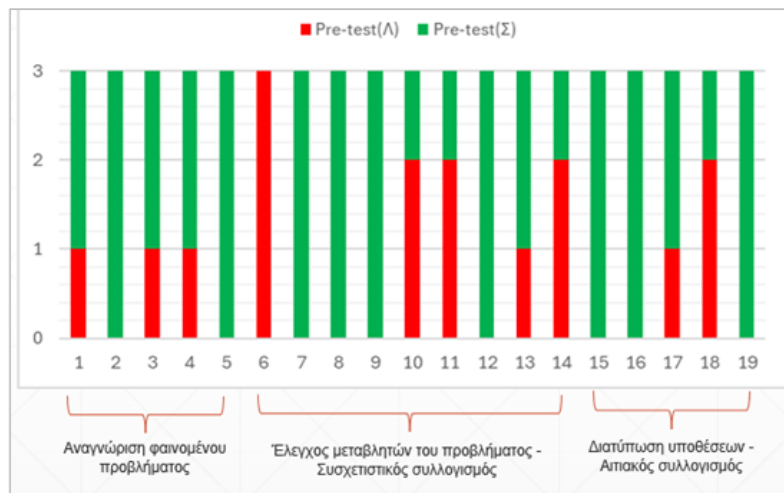
Αποτελέσματα

Στην παρούσα ενότητα παρουσιάζονται και αναλύονται τα αποτελέσματα της πιλοτικής έρευνας που διεξήχθη με τρεις μαθητές της Β' Λυκείου. Στο σχήμα 2 παρουσιάζονται οι σωστές και λάθος απαντήσεις των μαθητών στο σύνολο των ερωτήσεων που αφορούσαν τις 3 δεξιότητες προς διερεύνηση. Από το σύνολο των απαντήσεων φαίνεται πως το γνωστικό επίπεδο των μαθητών/-τριών είναι υψηλό.

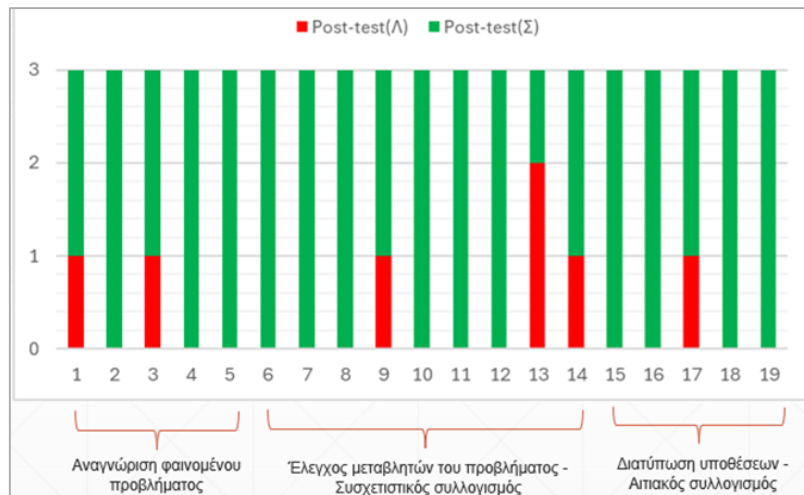
Η ποσοτική ανάλυση των αποτελεσμάτων του pre και post-test αναφορικά με τις τρεις διερευνώμενες δεξιότητες επιστημονικού συλλογισμού φανέρωσε συγκεκριμένες μεταβολές στις επιδόσεις των μαθητών (Σχήμα 2 και Σχήμα 3). Ως προς τη δεξιότητα της αναγνώρισης φαινομένου, σημειώθηκε μια μικρή αύξηση στις σωστές απαντήσεις, ανερχόμενες σε 13 στο post-test από 12 στο pre-test. Η αξιολόγηση της δεξιότητας του Ελέγχου Μεταβλητών και του Συσχετιστικού Συλλογισμού (Ερωτήσεις 6-14) ανέδειξε συνολικά θετικό πρόσημο βελτίωσης στις επιδόσεις των μαθητών μετά την παρέμβαση, όπου οι σωστές απαντήσεις αυξήθηκαν σημαντικά από 17 στο pre-test σε 23 στο post-test. Ιδιαίτερα ενδεικτική ήταν η πρόοδος στην Ερώτηση 6, η οποία διερευνούσε την αναγνώριση μεταβλητών στην οριζόντια βολή. Ενώ στο pre-test όλες οι απαντήσεις ήταν λανθασμένες, στο post-test παρατηρήθηκε πλήρης επιτυχία, υποδηλώνοντας έναν πιο διερευνητικό τρόπο προσέγγισης των μετρήσιμων μεγεθών που εμπλέκονται στο φαινόμενο από τους μαθητές μετά την παρέμβαση. Σε ό, τι αφορά την τρίτη δεξιότητα της διατύπωσης υποθέσεων και του αιτιακού συλλογισμού, παρατηρήθηκε επίσης αύξηση, με 14 σωστές απαντήσεις στο post-test έναντι 12 στο pre-test. Στην Ερώτηση 18, η οποία διερευνούσε την ορθή ερμηνεία της επίδρασης των δυνάμεων στην ελεύθερη πτώση, παρατηρήθηκε σημαντική βελτίωση. Ενώ στο pre-test μόνο 1 στους 3 μαθητές απάντησε σωστά, στο post-test και οι 3 μαθητές έδωσαν την ορθή απάντηση. Αυτό αποτελεί ένδειξη πως η αλληλεπίδραση σε θέματα σχέσης αιτίας-αποτελέσματος με το μοντέλο φαίνεται να βοήθησε τους μαθητές.

Τέλος, σε δύο ερωτήσεις παρατηρήθηκε μείωση των σωστών απαντήσεων. Στην Ερώτηση 9, που εξετάζει τη συσχέτιση μεταβλητών και συγκεκριμένα τις εξαρτήσεις της ταχύτητας, καταγράφηκε μία λανθασμένη απάντηση στο post-test, η οποία αποδόθηκε σε απροσεξία και όχι σε γνωστική δυσκολία, όπως ανέφερε η μαθήτρια. Στην Ερώτηση 13, που αφορά στις ενεργειακές μεταβολές στην οριζόντια βολή, οι σωστές απαντήσεις μειώθηκαν από δύο σε μία. Η εσφαλμένη επιλογή «η κινητική ενέργεια διατηρείται σταθερή, ενώ η δυναμική μειώνεται» οφείλεται σε παρερμηνεία, καθώς ο μαθητής εστίασε στη σταθερότητα της οριζόντιας συνιστώσας της ταχύτητας, αγνοώντας τη συνολική αύξηση της κινητικής ενέργειας λόγω της κατακόρυφης επιτάχυνσης. Συνολικά τα αποτελέσματα αποτύπωσαν βελτίωση των μαθητών στις τρεις δεξιότητες που διερευνήθηκαν.

Σχήμα 2. Οι απαντήσεις των μαθητών/-τριών στις ερωτήσεις γνωστικού αντικείμενου στο pre-test.



Σχήμα 3. Οι απαντήσεις των μαθητών/-τριών στις ερωτήσεις γνωστικού αντικείμενου στο post-test.



Πέρα από τις βελτιώσεις στις δεξιότητες επιστημονικού συλλογισμού, η έρευνα ανέδειξε μια ιδιαίτερα θετική εμπειρία των μαθητών από την αλληλεπίδραση με το LLM ως καθοδηγητή. Τα αποτελέσματα της αξιολόγησης, μέσω ερωτηματολογίου 13 ερωτήσεων 5-βάθμιας κλίμακας Likert (βλ. Παράρτημα), έδειξαν συντριπτικά θετική ανταπόκριση (κυρίως βαθμοί 4: Συμφωνώ και 5: Συμφωνώ απόλυτα) τόσο στη γενική εμπειρία (Ερωτήσεις 1-6) όσο και στην υποστήριξη της μάθησης (Ερωτήσεις 7-13). Οι μαθητές βρήκαν την καθοδήγηση σαφή, τις ερωτήσεις βοηθητικές για τη βήμα προς βήμα σκέψη, και την ανατροφοδότηση

χρήσιμη. Επιπλέον, εξέφρασαν την άποψη ότι η αλληλεπίδραση βελτίωσε την ικανότητά τους στην επίλυση προβλημάτων Φυσικής και δήλωσαν πρόθυμοι να χρησιμοποιήσουν ξανά παρόμοιο σύστημα. Αυτά τα ευρήματα ενισχύουν σημαντικά την αξιοπιστία και τη δυνητική χρησιμότητα της προτεινόμενης μεθοδολογίας καθοδήγησης από LLM.

Συζήτηση & Συμπεράσματα

Σε ό, τι αφορά το πρώτο ερευνητικό ερώτημα και την επίδραση της καθοδήγησης από LLM, η πιλοτική έρευνα έδειξε ότι η δομημένη αλληλεπίδραση με συγκεκριμένη αλυσίδα σκέψης σε συνδυασμό με την εφαρμογή τεχνικών Prompt Engineering (όπως η προσαρμοστική καθοδήγηση και η δυναμική ανατροφοδότηση), συνέβαλε θετικά στην ενίσχυση των δεξιοτήτων επιστημονικού συλλογισμού των μαθητών στην επίλυση προβλημάτων Φυσικής. Αυτό φάνηκε από τη συνολική βελτίωση στις επιδόσεις των μαθητών στο post-test και στις 3 αξιολογούμενες δεξιότητες (αναγνώριση φαινομένου, έλεγχος μεταβλητών/συσχετιστικός συλλογισμός, διατύπωση υποθέσεων/αιτιακός συλλογισμός).

Τα ευρήματα της παρούσας πιλοτικής μελέτης επιβεβαιώνουν τη δυναμική των Μεγάλων Γλωσσικών Μοντέλων (LLMs) στη βελτίωση των δεξιοτήτων επιστημονικού συλλογισμού των μαθητών στη Φυσική. Η καταγεγραμμένη πρόοδος σε τομείς όπως η αναγνώριση φαινομένων, ο έλεγχος μεταβλητών και η διατύπωση υποθέσεων, ευθυγραμμίζεται με τα συμπεράσματα των Jiang και Jiang (2024), οι οποίοι έδειξαν ότι η αλληλεπίδραση μαθητών με σύστημα υποστήριξης βασισμένο σε LLM ενισχύει τη βαθύτερη εννοιολογική κατανόηση και την ικανότητα διατύπωσης αιτιακών σχέσεων. Επιπλέον, η βελτίωση που καταγράφηκε κυρίως στον έλεγχο μεταβλητών και στον αιτιακό συλλογισμό συνάδει με προηγούμενες μελέτες (Chen & Klahr, 1999), οι οποίες τεκμηριώνουν ότι η καθοδηγούμενη παρέμβαση ενισχύει τη δυνατότητα των μαθητών να αναγνωρίζουν κρίσιμες μεταβλητές και να ερμηνεύουν σχέσεις αιτίας-αποτελέσματος.

Τέλος, η χρήση μεγάλων γλωσσικών μοντέλων (LLM) για την παροχή ανατροφοδότησης και υποστήριξης στη μάθηση, καθιστά αναγκαία την ανάπτυξη κριτικής σκέψης εκ μέρους των χρηστών. Παρόλο που τα LLM μπορούν να παράγουν πλούσιες και φαινομενικά έγκυρες πληροφορίες, είναι σημαντικό οι χρήστες να αξιολογούν κριτικά την ακρίβεια, τη συνάφεια και την εγκυρότητα των απαντήσεων, καθώς τα μοντέλα αυτά ενδέχεται να παράγουν περιεχόμενο που είναι ατελές, ανακριβές ή επηρεασμένο από περιορισμούς που αποδίδονται στα δεδομένα εκπαίδευσής τους (Bender et al., 2021). Παρότι τα πρώτα ευρήματα της παρούσας πιλοτικής μελέτης κατέδειξαν θετική επίδραση της παρέμβασης στη μαθησιακή εμπειρία, πρέπει να αναγνωριστεί ότι το μικρό και περιορισμένο δείγμα δεν επιτρέπει τη γενίκευση των αποτελεσμάτων. Αντίθετα, η μελέτη είχε ως στόχο μια πρώτη διερεύνηση τρόπων ενσωμάτωσης τέτοιων συστημάτων στη μαθησιακή διαδικασία, με σκοπό να εφαρμοστούν παρόμοιες παρεμβάσεις σε μεγαλύτερο δείγμα και πιο συστηματικά στο μέλλον.

Βιβλιογραφία

- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big?. Στο *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*, σσ. 610-623. <https://doi.org/10.1145/3442188.3445922>
- Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., ... & Zhang, Y. (2023). Sparks of artificial general intelligence: Early experiments with gpt-4. *arXiv preprint arXiv:2303.12712*. <https://doi.org/10.48550/arXiv.2303.12712>
- Chen, Z., & Klahr, D. (1999). All other things being equal: Acquisition and transfer of the control of variables strategy. *Child development*, 70(5), 1098-1120. <https://doi.org/10.1111/1467-8624.00081>
- Chi, M. T., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive science*, 5(2), 121-152. <https://www.sciencedirect.com/science/article/pii/S0364021381800298>

- Jiang, Z., & Jiang, M. (2024). Beyond answers: Large language model-powered tutoring system in physics education for deep learning and precise understanding. *arXiv preprint* arXiv:2406.10934. <https://doi.org/10.48550/arXiv.2406.10934>
- Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive science*, 11(1), 65-100. [https://doi.org/10.1016/S0364-0213\(87\)80026-5](https://doi.org/10.1016/S0364-0213(87)80026-5)
- Lawson, A. E. (1978). The development and validation of a classroom test of formal reasoning. *Journal of Research in Science Teaching*, 15(1), 11-24. <https://doi.org/10.1002/tea.3660150103>
- Lawson, A. E. (2000). The generality of hypothetico-deductive reasoning: Making scientific thinking explicit. *The American Biology Teacher*, 62(7), 482-495. <https://doi.org/10.2307/4450956>
- MacIsaac, D. (2023). Chatbots attempt physics homework—chatgpt: Chat generative pre-trained transformer. *The Physics Teacher*, 61(4), 318-318. <https://doi.org/10.1119/10.0017700>
- Polya, G. (2014). How to solve it: A new aspect of mathematical method. Princeton university press. ISBN: 9780691164076
- Sahoo, P., Singh, A. K., Saha, S., Jain, V., Mondal, S., & Chadha, A. (2024). A systematic survey of prompt engineering in large language models: Techniques and applications. *arXiv preprint* arXiv:2402.07927. <https://doi.org/10.48550/arXiv.2402.07927>
- Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., ... & Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems*, 35, 24824-24837. https://proceedings.neurips.cc/paper_files/paper/2022/file/9d5609613524ecf4f15af0f7b31abca4-Paper-Conference.pdf
- White, J., Hays, S., Fu, Q., Spencer-Smith, J., & Schmidt, D. C. (2024). Chatgpt prompt patterns for improving code quality, refactoring, requirements elicitation, and software design. In *Generative AI for Effective Software Development*, σσ. 71-108. Springer Nature Switzerland. <https://doi.org/10.1007/978-3-031-55642-5>
- Yeadon, W., Inyang, O. O., Mizouri, A., Peach, A., & Testrow, C. P. (2023). The death of the short-form physics essay in the coming AI revolution. *Physics Education*, 58(3), 035027. <https://doi.org/10.48550/arXiv.2212.11661>
- Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. *Developmental review*, 27(2), 172-223. <https://doi.org/10.1016/j.dr.2006.12.001>

Παράρτημα

Ερωτηματολόγιο 1 (Pre-Post test):

https://drive.google.com/file/d/1M1SCcGA4QX7xmLZDWAb_He_8OxD9v_NY/view?usp=sharing

Ερωτηματολόγιο 2 (Αξιολόγηση μαθησιακής εμπειρίας):

<https://drive.google.com/file/d/1RZqSy7rABNJBGsF4ZR0MNwEznVQsOLB7/view?usp=sharing>