

אפיון שיחות מתמטיות של פרחי הוראה עם מודלי שפה גדולים (LLMs) באמצעות עדשת הוראה רספונסיבית: הדוגמה של חילוץ מטעות

ד"ר רז הראל

בעל תואר שלישי בחינוך מתמטי מאוניברסיטת תל אביב ובוגר לימודי בתר־דוקטורט באוניברסיטת חיפה. מכהן כראש החוג למתמטיקה, מרצה ומדריך פדגוגי במכללה האקדמית לחינוך ע"ש דוד ילון.



ד"ר נדב מרקו

בעל תואר שלישי בחינוך מתמטי מהאוניברסיטה העברית בירושלים. מרצה ומדריך פדגוגי במכללה האקדמית לחינוך דוד ילון, ועוסק בהכשרת מורים למתמטיקה ובהוראת קורסים במתמטיקה ובחינוך מתמטי. חוקר בתר־דוקטורט בבית הספר לחינוך ע"ש סימור פוקס באוניברסיטה העברית בירושלים, במרכז לחקר למידת מורים והתפתחותם.



ד"ר רעות פרשה

בעלת תואר שלישי בחינוך מתמטי ממכון ויצמן למדע. ראש החוג למתמטיקה ומרצה במכללה האקדמית לחינוך אורות ישראל, ועוסקת בהכשרת מורים למתמטיקה ובהוראת קורסים במתמטיקה ובחינוך מתמטי. יועצת אקדמית, מנחת קהילות מורים וחברה בקבוצות המחקר של פט"ל למתמטיקה, העוסק בעיצוב משימות מקדמות הבנה, ושל פרויקט מס"ר העוסק בעיצוב סימולציות הוראה מבוססות AI במכון ויצמן למדע.



ד"ר גיל שורץ

מרצה בכירה במחלקה להוראת המדעים במכון ויצמן למדע, בקבוצת המתמטיקה, וראש מעבדת Simulating Mathematically Responsive Teaching. היא מקדמת מחקר ופיתוח של סימולציות דיגיטליות מבוססות אוטארים ובינה מלאכותית ומנהלת את פרויקט מס"ר (מעצבים סימולציות רספונסיביות).



אפיון שיחות מתמטיות של פרחי הוראה עם מודלי שפה גדולים (LLMs) באמצעות עדשת הוראה רספונסיבית: הדוגמה של חילוף מטעות

רז הראל, נדב מרקו, רעות פרשה וגיל שורץ

תקציר

האימוץ המהיר של מודלי שפה גדולים (LLMs) במרחבים חינוכיים יוצר מציאות חדשה שבה מורים ופרחי הוראה מנהלים דרישה מתמטי ופדגוגי עם מודלים שפתיים אלו, הזמינים לכל משתמש. על אף הפוטנציאל הפדגוגי שנקשר לכלים אלו, טרם התגבשה הבנה מחקרית שיטתית על יכולתם לתמוך בפיתוח של פרקטיקות ליבה של הוראת מתמטיקה ועקרונות העיצוב הנדרשים לפיתוח זה. במחקר הזה נתמקד באינטראקציות בין פרחי הוראה למתמטיקה ובין מודלי שפה גדולים (LLM) באחת הפרקטיקות המאתגרות בהוראה: חילוף תלמידים מטעויות. באינטראקציה לימודית עם תלמידים, ועל פי העקרונות של הוראה רספונסיבית מתמטית (MRT – Mathematically Responsive Teaching), פרקטיקה זו דורשת מהמורה לא רק לזהות טעויות ולדון בהן במפורש אלא גם לראות בנקודת המבט המתמטית של התלמיד משאב לחילוף מהטעות. במחקר זה אנו מבקשים לבדוק כיצד פרחי הוראה מחלצים מודלי שפה מטעויות בעת פעילות בקורס שבה התבקשו לאתגר מודלים עם בעיה מתמטית, ולאחר מכן לחלץ אותם מהטעות, כלומר להדריך מודל שפה גדול כיצד להגיע לתשובה הנכונה. לשם כך ניתחנו 20 שיחות בין פרחי הוראה למודלי שפה בעזרת כלי ניתוח שיח, ובמאמר זה אנו מתארים את המאפיינים הרספונסיביים בשיחות אלו. לניתוח האינטראקציות נעזרנו במסגרת תאורטית לשונית, SFL (Systemic Functional Linguistics) שאפשרה לנו להגדיר רספונסיביות מתמטית מבחינה אופרטיבית. מהממצאים עולה שיש הבדל ניכר במידת הרספונסיביות בין פרחי ההוראה, ורספונסיביות יכולה להתפתח בדינמיות בתוך שיחה אחת. תרומת המחקר היא כפולה: 1. מתודולוגית-תאורטית – פיתוח הכלי SFL-MRT שישמש לניתוח רספונסיביות מתמטית באינטראקציות עם מערכות בינה מלאכותית (AI); 2. פדגוגית – הראייה שדרישה עם מודלי שפה גדולים בתנאים עיצוביים מסוימים עשויה לשמש פלטפורמה פורייה לפיתוח הוראה רספונסיבית.

מילות מפתח: מודלי שפה גדולים (LLMs); הוראה רספונסיבית מתמטית; חילוף מטעויות; הכשרת מורים; SFL; בינה מלאכותית.

מבוא

התפתחויות אחרונות במודלי שפה גדולים (LLMs) מאפשרות דיאלוג בין משתמשים אנושיים למערכות בינה מלאכותית (AI) במגוון רחב של תחומים, בהם הוראת מתמטיקה (Taani & Alabidi, 2024). בעוד מודלי שפה מסוגלים לייצר שיח מתמטי שוטף, תשובותיהם מכילות לעיתים חוסר דיוק ושגיאות מתמטיות המוצגות בביטחון הציעו לבחון טעויות מתמטיות אלו לא כמגבלות של הכלי אלא כבעלות פוטנציאל פדגוגי לתרגול של פעילויות מתמטיות מקוריות חשובות, כגון העלאת בעיות ("problem posing"; Silver, 1997), הבנת הוכחות מתמטיות ("proof comprehension"; Mejía-Ramos et al., 2012; Stylianides & Stylianides, 2024), תיקוף הוכחות ("proof validation"; Selden & Selden, 2003) ותרגול חשיבה ביקורתית (Cassinadri, 2024). במחקר הנוכחי אנו מבקשים לבחון יישום של פוטנציאל פדגוגי זה בתחום של הכשרת מורים. הנחת המחקר היא שמודלי שפה גדולים מדמים תקשורת אנושית ברמה המאפשרת לפרחי הוראה לא רק לזהות טעויות מתמטיות שנעשו, אלא גם להפעיל מיומנויות ליבה פדגוגיות בזמן שהם מסייעים בתיקון הטעויות.

חילוף תלמידים מטעות מתמטית בדרך פורה ומקדמת הוא אחת הפרקטיקות המאתגרות בהוראה שדורשת מהמורה לזהות את הטעות, להבין את ההיסק המתמטי שהוביל אליה ואת ההיגיון שבו, ולראות בהם משאב ללמידה. כל זאת מתוך ניגוד לנטייה להתעלם או לבטל את הטעות ולהציג דרך פתרון חלופית במקומה (Borasi, 1994). מחקרים בתחום החינוך המתמטי רואים בחילוף מטעות פרקטיקה פדגוגית מורכבת ורצויה: ברודי (Brodie, 2014) רואה בטעות מתמטית תשובה של הלומד אשר משקפת את דרך החשיבה הנוכחית שלו ואת ההיגיון הפנימי שמאחוריה. על פי הבחנה זו, חילוף מטעות מצריך מהמורה לעבור שלושה שלבים: זיהוי הטעות, פרשנותה – כלומר הבנת ההיגיון שהוביל אליה – ועיסוק פעיל בה כך שתשמש משאב לפיתוח החשיבה המתמטית. לפי בריי (Bray, 2011), מורים הנוטים לעסוק בטעויות ולא לתקן אותן



ישירות מצליחים לעורר שיח מתמטי עמוק יותר בכיתה. ממצאים אלו תואמים את העמדה שלפיה חילוף מטעות אינו שם נרדף לתיקון הטעות, אלא משמש פרקטיקה המצריכה העמקה בחשיבה שהובילה לטעות, מציאת ההיגיון העומד בבסיסה ובנייה על היגיון זה.

העיסוק בחילוף מטעויות הוא נושא עיקרי במחקר החינוכי העוסק בהוראה רספונסיבית מתמטית (MRT – Mathematically Responsive Teaching) (Bishop et al., 2022; Robertson et al., 2016; Schwartz, 2025). גישה זו מדגישה את היחס של המורים ותגובותיהם כלפי הרעיונות של תלמידיהם, נקודות המבט שלהם וחוויותיהם בעת אינטראקציה מתמטית. בספרות של הוראה רספונסיבית, תגובה פרודוקטיבית לשגיאות כוללת העמקה בחשיבה המתמטית שהובילה לטעות, מתוך תפיסה של הטעות כמשאב ללמידה (Robertson et al., 2016). במחקר זה, העדשה התאורטית הנ"ל תשמש אותנו לבחינת דו־שיח בין פרחי הוראה למודלים שפתיים גדולים.

שיחות מתמטיות בין מורים למודלי שפה הן תופעה חדשה יחסית המתפתחת במהירות. אחד הכיוונים הבולטים בתחום זה הוא תחום המחקר המתפתח של סימולציות הוראה מבוססות בינה מלאכותית, שבהן בוטים המתוכנתים לְדַמּוּת תלמידים מציגים טעויות נפוצות ומאפשרים למורים ולפרחי הוראה לתרגל פרקטיקות הוראה מגוונות, בהן חילוף חשיבה מתמטית (Schwartz & Czaczkes, 2026; Zhuang & Zhang, 2025).

ואכן, השימוש במודלי שפה גדולים לתמיכה בחילוף מטעות זוכה לבחינה מחקרית עדכנית. ואנג ואחרים (Wang et al., 2024) פיתחו מסגרת שבה מודלי שפה מסייעים למורים-מדריכים בתהליך של אבחון סוג הטעות, בחירת אסטרטגיה וניסוח תגובה. ממצאיהם עולה כי מודלים אלו מסוגלים לשפר את איכות ההתמודדות של התלמידים עם טעויות. לין ואחרים (Lin et al., 2025) הדגישו כיצד ניתוח טעויות של מודלי שפה גדולים שעושים תלמידים בבית ספר יסודי תורם לפיתוח חשיבה מטא-קוגניטיבית ומעמיק את ההבנה המושגית שלהם. מחקרים אלו מלמדים על פוטנציאל פדגוגי של מודלי שפה גדולים בלמידה מטעויות, אך הם מתמקדים בתהליכי אבחון ותיקון של טעויות, ולא בחילוף רספונסיבי הנשען על חשיבת הלומדים, שהוא מוקד המחקר הנוכחי.

מחקר זה מבקש לתרום לתחום המתפתח באמצעות בחינת הדרכים שבהן פרחי הוראה בוחנים ומגיבים על השגיאות המתמטיות של מודל השפה באמצעות עדשת הוראה רספונסיבית למתמטיקה. ניתוח זה הוא צעד ראשון לעבר הבנה מחקרית רחבה יותר, השואפת לתמוך בפיתוח רספונסיביות בקרב פרחי הוראה המתמטיקה בעזרת התערבויות מגוונות. בפרט אנו שואלים: אילו מאפיינים של הוראה רספונסיבית למתמטיקה באים לידי ביטוי בשיחות של פרחי הוראה עם מודלי שפה גדולים בעת ניסיונות לחלץ את המודל מטעות מתמטית?

סקירת ספרות

הוראה רספונסיבית למתמטיקה

הוראה רספונסיבית היא גישת הוראה המעמידה במרכז את חשיבת התלמידים ואת השימוש ברעיונותיהם בתהליך ההוראה (Jacobs et al., 2010; Robertson et al., 2016). בחינוך מתמטי, הוראה רספונסיבית כוללת זיהוי של המהות המתמטית ברעיונות הלומדים, פרשנות שלהם ושילובם בתהליכי הוראה (Bishop et al., 2022). במחקר זה אנו מבקשים להרחיב את ההמשגה של הוראה רספונסיבית למתמטיקה מעבר למהלכי הוראה רספונסיביים המוכרים בספרות (e.g., Jacobs & Empson, 2016; Lee et al., 2020): על פי בובר (Buber, 1958), הוראה בשיח בין המורה לתלמיד מפגש המושתת על פתיחות, נוכחות והקפה של המורה את התלמיד, אנו רואים בהוראה רספונסיבית למתמטיקה הוראה שמדגישה את אחריות המורה לזהות את נקודת המבט המתמטית של הלומד בעת האינטראקציה, ולראות בה בסיס לדיון המתמטי ביניהם (Robertson et al., 2016; Schwartz, 2025).

אחת הפרקטיקות של הוראה רספונסיבית, חילוף תלמידים מטעויות מתמטיות בדרך שנשענת על ההיגיון שלהם, היא פרקטיקה הכרחית ומאתגרת במיוחד: היא דורשת מהמורה לאמץ את נקודת המבט של התלמיד, להבין את דרכי המחשבה שלו שהובילו לטעות, ולזהות רכיבים נכונים או הגיוניים בדבריו כדי לבנות עליהם. התעלמות מרעיונות תלמידים או התמקדות בהסבר של התשובה הנכונה בלבד, עלולות למנוע מתלמידים לפתח הבנה מושגית שמחוברת לידע הקודם שלהם. לפיכך חילוף מטעות משמש קרקע פורייה לבחינת רספונסיביות מתמטית בהוראה. במחקר הזה אנו מרחיבים גישה זו ומיישמים אותה בשיח עם מודל השפה. אנו גורסים כי זיהוי המרכיבים ההגיוניים בטעות מתמטית של מודל שפה, והצורך לגשר בינם ובין המושג המתמטי הנכון, מחייבים את המורה לפתח סקרנות כלפי נקודת המבט של האחר, בדומה לאינטראקציה של מורה עם תלמיד אמיתי ששוגה בשיעור. מתוך כך, השיחה עם מודל שפה גדול הופכת לזירה שבה אפשר לאפיין ולבחון את התממשותה של רספונסיביות מתמטית, גם אם בין מורה למודל שפה.

כדי לזהות מופעים של רספונסיביות מתמטית באינטראקציות בין מורים למודלי שפה גדולים, אנו נשענים על כלי ניתוח שיח המסתמכים על SFL (Systemic Functional Linguistics; Halliday & Matthiessen, 2004). זוהי תאוריה בלשנית המקובלת בתחומי מחקר רבים, ובייחוד בחינוך, המאפשרת לנתח כיצד משמעות נבנית באמצעות שפה. לפי SFL, השפה נחשבת לא רק אמצעי להעברת מידע, אלא משאב לניהול יחסים בין משתתפים ולארגון הידע המובע בשיח, ולפיכך אפשר לנתח כל אמירה בשיח בשלושה ממדים: 1. במה האמירה עוסקת – איזה רעיון או התנסות בעולם היא מבטאת (הממד הרעיוני, ideational); 2. איך היא עוסקת ברעיון – אילו מערכות יחסים היא מבנה (הממד הבין-אישי, interpersonal); 3. איך האמירה משתלבת במארג שיח – איך המסר מתארגן בתוך השיחה (הממד הטקסטואלי, textual). הבחירה במסגרת תאורטית זו ככלי לניתוח מוצדקת מכיוון שאנו רואים בהוראה רספונסיביות הוראה שמתבטאת לא רק בכוונות פדגוגיות או בדיונים מסוימים, כמו שמציעים מאמרים רבים בתחום (ראו Jacobs & Empson, 2016), אלא גם בשימוש המפורש של מורים בשפה: בבחירה שלהם האם להעדיף שאלות או אמירות, בוודאות שבה הם מציגים טענות מתמטיות, ובמידה שבה הם מנהלים שיחה שיש בה מבעים רבים הנוגעים במישרין לרעיונות המתמטיים של שותפם לשיחה.

ב-SFL השפה נחשבת מערכת (system) של בחירות (למשל, כאשר דובר רוצה להביע רעיון מסוים, הדובר בוחר בשאלה, הצהרה או פקודה). בחירות אלו תלויות בהקשר התקשורתי, והן אלו שמעצבות את המשמעות המועברת. מבחינה מתודולוגית, אפשר לראות במערכות קטגוריות המארגנות את הרקע התאורטי של המאמר. במאמר הזה אנו מתמקדים בשלוש מערכות: שתיים הקשורות לממד הבין-אישי ואחת קשורה לממד הטקסטואלי. בחרנו במערכות אלו משום שהן מרכזיות בניתוח שיח חינוכי ושימשו בעבר בחקר חינוך מתמטי (למשל, Ebbelind, 2023), ומשום שהן רלוונטיות במיוחד לדיון ברספונסיביות מתמטית. בשל אופי הנתונים, במחקר זה בחרנו שלא לנתח בעזרת הממד הרעיוני, אך הוא בהחלט יהיה חלק מהתחום שאנו מציעים במחקרי המשך. נציג להלן את המערכות המרכזיות שנשתמש בהן לניתוח הנתונים.

מערכת ה־Mood

בממד הבין-אישי, מערכת ה־Mood עוזרת להבחין כיצד מורים ממצבים את עצמם בשיחה: כסמכות בעלת ידע או דווקא כמי שמקדמים שיח וחקירה עם האחר. המערכת מבחינה בין סוגי פניות שדובר עשוי לבחור בהן בשיחה, כגון שאלה פתוחה, שאלה סגורה, קביעה, או הוראה/בקשה, כלומר עוזרת לשקף את התפקיד שהמורה ממלא באינטראקציה – האם הוא מבקש מידע, מוסר מידע, מנחה את בן השיח או שותף לחשיבה משותפת. מתוך כך אפשר לראות אם המורה ממצב את עצמו כסמכות בעלת ידע, כמקדם דיאלוג, או שילוב כלשהו של גישות אלו.

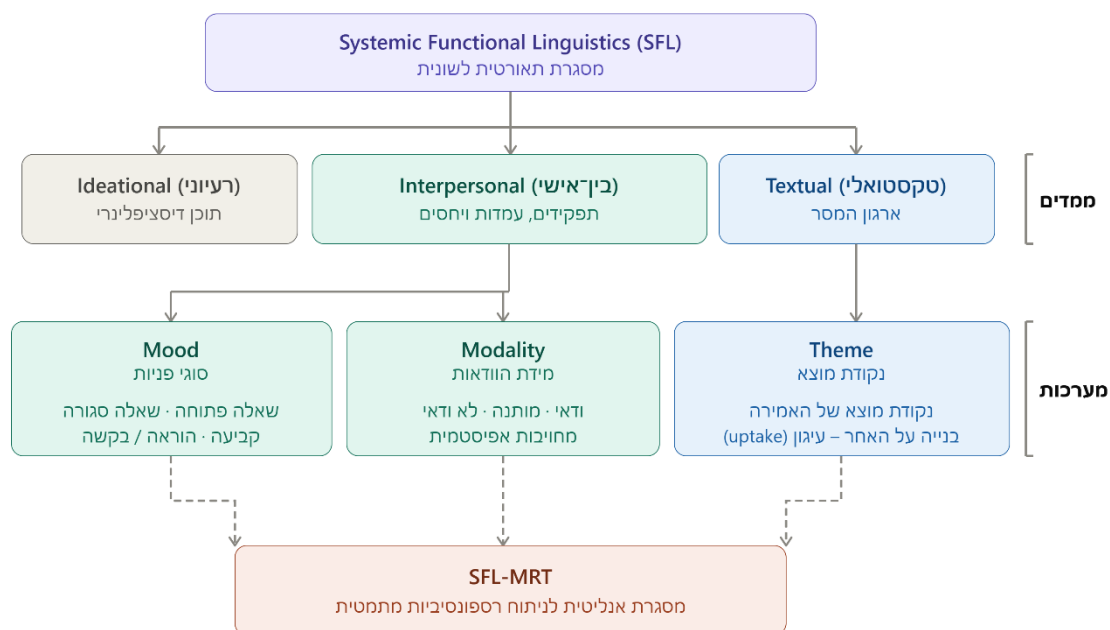
מערכת המודליות (Modality)

גם בממד הבין-אישי, מערכת המודליות מבחינה בין בחירות לשוניות הקשורות למידת הוודאות, המחויבות או הפתיחות של הדובר כלפי הטענה. האופן שבו טענות מתמטיות מוצגות – כעובדה ודאית, כהשערה, כמשפט פתוח לדיון, או כטענה הדורשת נימוק – מלמד עד כמה הדובר משאיר לבן השיח מרחב להשתתף, לשאול ולהציע רעיונות. שילובן של שתי המערכות (Mood ו־Modality) מאפשר להבין כיצד פרחי ההוראה מגיבים כלפי בן שיח בעת טעות וכיצד הם מאזנים בין הצגת ידע מתמטי מדויק ובין עידוד השתתפות פעילה ושיח פתוח.

מערכת התמה (Theme)

בממד הטקסטואלי, מערכת התמה עוסקת בבחירת הדוברים לגבי מוקד המסר שלהם ובאיזה אופן הוא קשור למוקד של המשפט הקודם, שאמר בן השיח. באמצעותה אפשר לבחון כיצד פרחי ההוראה קושרים את דבריהם למה שנאמר קודם בשיח, ובייחוד האם הם מסתמכים על האמירה שהציג מודל השפה: האם הם ממשיכים, מפתחים או מתקנים אותה, או לחלופין מתעלמים ממנה ומציגים מהלך חדש. ניתוח זה מאפשר לבדוק האם ובאיזו מידה מתרחש עיגון (uptake) בחילופי הדברים, כלומר הישענות ישירה על דברי האחר.

שלוש המערכות יחד מאפשרות להתחקות אחר רספונסיביות ולראות בה תהליך דיאלוגי המתפתח במהלך השיחה. מסגרת משולבת זו תיקרא להלן SFL-MRT, מכיוון שבאמצעותה בחרנו מערכות מסוימות של SFL שמאפשרות לשפוך אור על מרכיבי MRT (Mathematically Responsive Teaching). המסגרת משמשת כלי אנליטי לבחינת האופן שבו רספונסיביות מתמטית מתעצבת ומתממשת באינטראקציות דבורות או כתובות. איור 1 מציג תרשים של המסגרת התאורטית שתיארנו עד כה.



איור 1: מסגרת אנליטית לניתוח רספונסיביות מתמטית SFL-MRT

מתודולוגיה

מחקר זה הוא מחקר גישוש איכותני הנתמך בניתוח כמותי. מטרתו לאפיין הוראה רספונסיבית למתמטיקה בקרב פרחי הוראה בעת אינטראקציות עם מודלי שפה גדולים (LLMs), ובייחוד בחילוף מטענות מתמטיות שהציג מודל השפה. בחנו האם וכיצד פרחי הוראה מנהלים דיאלוג מתמטי רספונסיבי עם מודלי השפה כאשר הם מנסים להביא את המודל לידי הכרה בטעותו ולתקן אותה. במחקר השתתפו 20 פרחי הוראה למתמטיקה הלומדים במכללה אקדמית לחינוך במרכז הארץ. בקבוצה נמצאים סטודנטים סדירים בשנת לימודיהם השנייה וסטודנטים בתוכנית להסבת אקדמאים בעלי רקע אקדמי קודם.

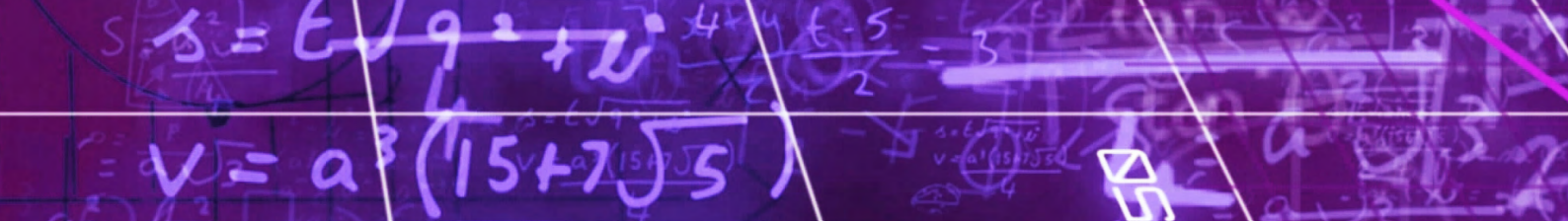
התקבלה הסכמה מדעת מכל המשתתפים ואת המחקר אישרה ועדת האתיקה המוסדית. שמות הסטודנטים מופיעים בעילום שם.

מהלך המחקר

כל משתתף ניהל אינטראקציה מתמטית עם מודל שפה לבחירתו (לדוגמה ChatGPT, Claude, Gemini). המשתתפים יכלו לבחור מודל שפה הזמין להם כדי להבטיח גישה מיידית ללא עלות ובלי צורך בהכשרה מוקדמת. תחילה, מרצה הקורס הדגים שיח רספונסיבי מכוון לטענות עם ChatGPT על פי המודל של מרקו וסטיליאנידס (Marco & Stylianides, 2025): המרצה הציג משימה גאומטרית שגרמה למודל השפתי לטעות, והדגים כיצד הוא עוזר לצ'אט לתקן את עצמו בעזרת רצף שאלות, ולא בעזרת הסבר ישיר. שנית, פרחי ההוראה התבקשו לעשות פעולה דומה, שבה הם מכוונים את מודל השפה לטעות, ולאחר מכן עוזרים לו לתקן את הטעות. לשם כך התבקשו פרחי ההוראה לבצע את השלבים האלה: 1. לזהות בעיה מתמטית מתוכנית הלימודים שעשויה לגרום למודל השפה לטעות; 2. לשוחח עם מודל השפה עד שיצליח להכיר בטעות ולתקנה; 3. (רשות) להציג בעיה דומה אך שונה במקצת כדי לבחון האם מודל השפה "למד" ומסוגל כעת ליישם את הפתרון גם מחוץ להקשר של הבעיה המקורית. הסטודנטים התבקשו להגיש את האינטראקציה במלואה (כולל טקסט ותמונות).

איסוף ניתוח של הממצאים

נאספו 20 שיחות שניהלו 20 פרחי הוראה עם מודלים שפתיים גדולים מגוונים. לצורך ניתוח של מידת הרספונסיביות של פרחי ההוראה, השתמשנו במסגרת SFL-MRT (ראו איור 1). בכל שיחה נותחו המבעים של המורים בלבד, שכן מטרת



המחקר הייתה לבחון את הרספונסיביות של פרחי ההוראה ולא של המודל השפתי. בהתאם לשיטת הניתוח המקובלת במסגרת ה-SFL, כלל המבעים של פרחי ההוראה חולקו לפסוקיות (clauses) – יחידות ניתוח בסיסיות המוגדרות משפטים הכוללים נושא ונשוא. תהליך זה הניב בסך הכול 144 פסוקיות לניתוח.

הנתונים נותחו בשני שלבים. בשלב הראשון נעשה ניתוח ברמת המיקרו, שבו כל פסוקית קודדה בנפרד באמצעות שלוש המערכות שהוזכרו לעיל. בשלב השני נעשה ניתוח ברמת המאקרו, שבו נבחנו הדפוסים והמגמות שעלו מכלל הפסוקיות. סכמת הקידוד פותחה באופן איטרטיבי.

בחלק הקודם הצגנו בפירוט את המסגרת התאורטית שבה בחרנו. כעת נפרט את הקידוד שעשינו בעזרתה (ראו טבלה 1). במערכת ה-Mood השתמשנו בארבעה קודים שמסווגים את דרך הפנייה של המורה למודל שפה גדול: 1. שאלה פתוחה (open interrogative) המזמינה הנמקה או הרחבה; 2. שאלה סגורה (closed interrogative), כלומר שאלה המבקשת מידע מסוים או תשובה מוגדרת; 3. הצהרה (declarative); 4. ציווי/הוראה (imperative), כלומר הוראה או בקשה ישירה לפעולה (למשל, "תחשב את השטח").

במערכת המודליות קידדנו לפי שתי רמות: גבוהה (high epistemic modality) ונמוכה (low/median epistemic modality). השתמשנו בקוד של הרמה הגבוהה כדי לקודד טענות שמוצגות כוודאיות או ככאלה שאינן פתוחות לערעור. לעומת זאת הקוד של הרמה הנמוכה/בינונית יושם בעבור טענות הנחשבות כהשערה, הצעה או ככאלה המזמינות עוד בדיקה, חשיבה מחדש או הנמקה.

בעזרת קידוד משולב של שתי המערכות אפיינו את רמת הפתיחות של כל פסוקית – האם היא פתוחה או סגורה. התחלנו עם הקידוד של מערכת Mood: שאלה פתוחה קודדה תמיד למבע פתוח, ואילו הצהרה וציווי קודדו תמיד למבע סגור. מערכת המודליות שימשה כלי משני כדי להכריע אם פסוקיות שקודדו כשאלה סגורה (closed interrogative) מהוות מבע פתוח או סגור. קודדנו שאלות סגורות למבעים פתוחים בפעמים שבהן רמת הוודאות שלהן הייתה נמוכה (medium/low epistemic modality). לדוגמה, בשאלה "האם אתה עדיין בטוח שבסיסי המנסרה הם משולשים?", לכאורה מדובר בשאלה סגורה, אך ההתמקדות ברמת הביטחון של בן השיח ("עדיין בטוח") מערערת את הוודאות של הטענה הקודמת ומזמינה בחינה מחודשת שלה. לכן אף שמבחינה צורנית זו שאלה סגורה, היא פותחת מרחב לחשיבה ולנימוק ומזמנת בחינה מחדש של אמירה קודמת, ועל כן קודדה למבע פתוח.

בממד הטקסטואלי, במערכת התמה, קודדנו את האופן שבו פרחי ההוראה בחנו את הרעיונות המתמטיים שהציג מודל השפה ובייחוד האם וכיצד הם בנו את דבריהם על בסיס נושא השיח המתמטי שהוצג קודם. בחינות אלו כונו עיגון, והן משקפות את מידת ההמשכיות, הקישור וההשתלבות של תרומת פרחי ההוראה במארג המתפתח של השיחה. הגדרנו קודים בשלוש רמות של עיגון (theme uptake):

1. היעדר עיגון: אין כל אזכור בדברי פרחי ההוראה לדרך החשיבה המתמטית של מודל שפה גדול, והתרומה שלו אינה משתלבת ברצף הרעיוני שנבנה קודם; 2. עיגון חלקי: יש אזכור מסוים לדברי מודל שפה גדול אך הוא שטחי או חלקי, ואינו יוצר המשכיות מלאה של קו החשיבה; 3. עיגון ישיר: יש אזכור מפורש וישיר לחשיבה המתמטית של מודל השפה, באופן שממשיך, מפתח או מבקר אותה ותורם ללכידות ולהתפתחות הרציפה של השיח המתמטי. בניתוח הכמותי איחדנו בין הקודים של "עיגון חלקי" ו"עיגון ישיר", כפי שמופיע בעמודה השמאלית של טבלה 1, המרכזת את הקשרים בין ההיבטים, הקטגוריות והקידודים לעיל.

טבלה 1: היבטים וקטגוריות לניתוח הפסוקיות

ממד	מערכת	קידוד	סיווג הפסוקיות לצורך הניתוח הכמותי
בין-אישי	Mood	ציווי	סגורה
		הצהרה	סגורה
		שאלה פתוחה	פתוחה
	מודליות	שאלה סגורה	פתוחה עבור מודליות נמוכה; סגורה עבור מודליות גבוהה
		גבוהה	*
טקסטואלי	עיגון (Theme Uptake)	נמוכה	*
		היעדר עיגון	ללא עיגון
		עיגון חלקי	מעוגנת
		עיגון ישיר	מעוגנת

* הערה לטבלה 1: בסיווג הרספונסיביות של הפסוקית לצורך הניתוח הכמותי, המודליות שימשה אותנו רק כדי להבחין האם שאלות רטוריות הן "פתוחות" או "סגורות".



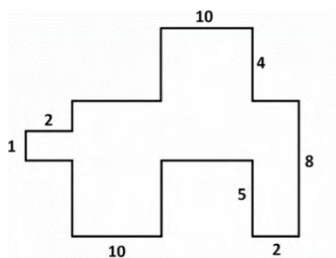
כדי להבטיח את מהימנות הקידוד, התחלנו בשלב כיוול, שבו נבחרו באקראי חמש שיחות שקידדו אותן שלושה מהמחברים בנפרד. בחינת הקידודים העלתה רמת הסכמה גבוהה מאוד בכל הקטגוריות, ובפעמים מעטות של חוסר עקיבות התנהל דיון משותף עד להגעה להסכמה מלאה. תהליך זה אישש את בהירותן של הגדרות הקידוד ואת יכולת היישום העקיבה שלהן. לאחר שלב הכיוול חולקו שאר האירועים באקראי בין המחברים להמשך קידוד עצמי. כדי לצמצם היתכנות של הטיות הנובעות מריבוי תפקידיו של המרצה במחקר זה, תהליך הקידוד נעשה מתוך בדיקה מתמדת עם שאר חוקרי הצוות, ונשען במשותף על קריטריוני הקידוד של מסגרת ה-SFL-MRT.

ממצאים

הממצאים של מחקר הגישוש יאורגנו להלן בדרך זו: תחילה נציג ניתוח של אירוע כדי לתאר בהרחבה כיצד ניתחנו את הרספונסיביות המתמטית או היעדרה בשיח בין פרחי ההוראה למודל השפתי, וכדי להציג לקוראים דוגמה מפורטת של אינטראקציה כזו. לאחר מכן נציג את ניתוח הרחוב של כלל האירועים. הדוגמה שתוצג להלן נבחרה בשל העושר היחסי שבה מבחינת סוגי מבעים, בהם הדגמה של שינוי ברמות הרספונסיביות בעת השיחה.

האינטראקציה של פרח ההוראה נטע עם הצ'אט

תחילה, נתאר את השיחה בין פרח ההוראה נטע (שם בדוי) ובין Chatgpt. איור 2 מציג את ההנחיה הראשונה שנתנה נטע לצ'אט ובו היא ביקשה ממנו לחשב את שטח הצורה המופיעה באיור.



"חשב את שטח הצורה המצורפת (התעלם מקנה המידה האמיתי והתייחס רק למידות הרשומות על מקטעי המצולע. אל תמדוד בכלל התייחס רק לצורה באופן כללי כמו שהיא נתונה)"

איור 2: השאלה שנטע הציגה כדי לגרום לצ'אט לעשות טעות מתמטית

להלן השורות העיקריות מתשובתו של מודל שפה גדול:

1. **נבנה מלבן חיצוני** שמקיף את כל הצורה:

רוחב = 10 (בר עליון) + 2 (הבליטה הימנית בתחתית) + 2 (השלב השמאלי העליון) ≈ 14

גובה = 4 (שלב עליון) + 8 (החלק הימני התחתון) = 12

שטח מלבן חיצוני $168 = 12 \times 14$

2. **נחסר את ה"חריצים"/הגומות** שמופיעות בתוך המלבן:

- גומה פנימית** (החור/הגומה התחתית במרכז) — גובהה = 5, רוחבה = $10 - 14 = -4$

שטח גומה $\approx 20 = 5 \times 4$

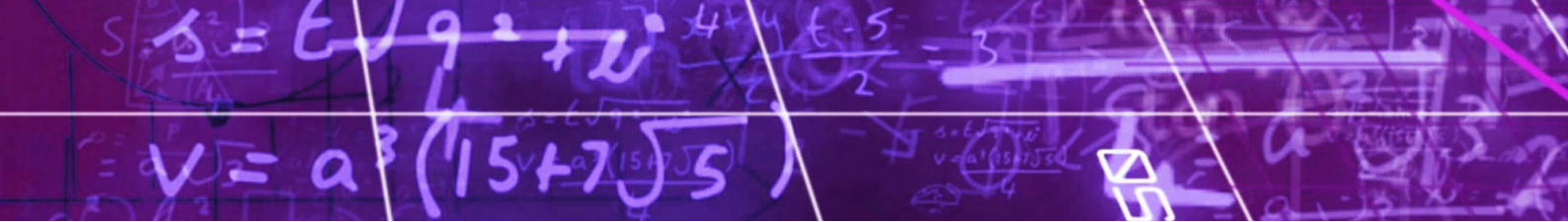
- בליטה שמאלית קטנה** — רוחבה ≈ 1 , גובהה ≈ 3

שטח בליטה $\approx 3 = 3 \times 1$

הערה: זהו תיקון מקומי של הפריסה כדי להתאים את כל המספרים.

3. **שטח המצולע הסופי:**

שטח $145 = 3 - 20 - 168$



אסטרטגיית החישוב של הצ'אט, שבה הוא הציע להקיף את הצורה במלבן ולחסר שטחים, הייתה נכונה. ואולם התוצאה הסופית שהגיע אליה הייתה שגויה (כלומר נטע הצליחה בחלק 1 של המשימה שקיבלה – לגרום למודל השפה לטעות מתמטית). שטח הצורה שהופיע בשאלה של נטע הוא 168 יחידות שטח, ולא 145 כפי שטען המודל השפתי. מעניין לשים לב כי המספר 168 אכן הופיע בשורה 3 בדברי הצ'אט, אך הוא לא הוצג כתוצאה סופית אלא כחלק מהחישובים (השגויים, שכן צלעות המלבן הן 24 ו-12). על כך הגיבה נטע: "איך מחשבים שטח צורה המורכבת ממלבנים?"

מבחינת ה-Mood קודדנו פסוקית זו לשאלה פתוחה (open interrogative) – שאלה המזמינה הסבר של הצ'אט על אסטרטגיית החישוב שלו. לכן מארבע הדרכים של ה-Mood (צייוי, הצהרה, שאלה סגורה, שאלה פתוחה) פסוקית זו קודדה לשאלה פתוחה. בנוגע למודליות (עמדה אפיסטמית), ראינו שדבריה של נטע דורשים הסבר והזמנה לדיאלוג. לכן קודדנו פסוקית זו למודליות נמוכה/בינונית (low/median epistemic modality). בסך הכול קודדנו את המבע לפתוח. באשר לעיגון, יש לשים לב שנטע אכן השתמשה בחלק מתשובתו של הצ'אט, אך לא עסקה באסטרטגיה המסוימת שלו וגם לא איתרה את הטעות. לפיכך מתוך שלוש רמות העיגון (היעדר עיגון, עיגון חלקי, עיגון ישיר) קודדה פסוקית זו לעיגון חלקי.

הצ'אט הגיב על דבריה של נטע בהסבר כללי על האופן שמחשבים את שטח הצורה הגאומטרית באמצעות פירוק למלבנים. נטע הגיבה: "ענה על השאלה הקודמת בשימוש בפירוק למלבנים".

מבחינת Mood, דברים אלו קודדו לצייוי (imperative) של נטע על הצ'אט, ובהם היא הכתיבה לו את דרך הפתרון. כיוון שהמורה דיברה בנחרצות קודדנו רמת מודליות גבוהה (high epistemic modality). בדבריה של המורה לא היה שום אזכור להיסק המתמטי של הבינה המלאכותית, ועל כן קודדו מבע זה להיעדר עיגון.

הצ'אט ניסה לחשב עוד פעם אחת, אך שוב נתן את אותה תשובה שגויה. נטע הגיבה כך: "ביקשתי שתחשב שטח בשימוש בשיטה המקובלת של חלוקה למלבנים בודדים וחיבור שטחים".

מדבריה של נטע עולה שהיא נחשבת סמכות הקובעת את אסטרטגיית החישוב המתאימה. כמו הפסוקית הקודמת, גם פסוקית זו קודדה לצייוי מבחינת ה-Mood, המשקפת רמת מודליות גבוהה ובלי עיגון. אפשר לתאר תסריט חלופי שבו נטע הייתה מזהה את מקור הטעות בדברי הצ'אט, ומתוך כך ממקמת את עצמה בעמדה אפיסטמית של אבחון והכוונה, המאפשרת חילוף יעיל של הפתרון. בתסריט כזה, בחירותיה הלשוניות היו עשויות לכוון את בן השיח לבחינה מחודשת של הנחותיו ולתיקון מהלך החשיבה. עם זאת, בפועל, נטע מימשה יכולות אחרות בתוך המשאבים הבינאריים והטקסטואליים שעמדו לרשותה, ובחרה לפעול בדרך אחרת. לאחר עוד ניסיון כושל מצד הצ'אט לחשב את שטח הצורה, נטע הגיבה כך: "אתה משתמש בשיטה של חיסור שטחים ואני מבקשת להשתמש בשיטה של חיבור מלבנים פשוטים".

משפט זה מכיל שתי פסוקיות. בפסוקית הראשונה בחנה נטע לראשונה את אסטרטגיית הפעולה של הצ'אט, ולכן היא קודדה עם עיגון ישיר ועם הצהרה (declarative) ורמת מודליות גבוהה. בפסוקית השנייה המשיכה נטע לצוות בנחרצות על הצ'אט כיצד עליו לפעול: צייוי, עם רמת מודליות גבוהה והיעדר עיגון.

לאחר ניסיון לא מוצלח אחר של הצ'אט להגיע לידי תוצאה הנכונה, כתבה נטע את הדברים האלה: "התשובה אינה נכונה. חלק את הצורה למלבנים כך שבחיבור השטחים תקבל 168".

ניסיונותיה של נטע לכוון את הצ'אט לדרך מסוימת וחוסר ההצלחה של ניסיונות אלו גרמו לה לצמצם עוד יותר את מרחב העשייה וקבלת ההחלטות של הצ'אט והיא מיקדה אותו ביעד מספרי מסוים, כלומר 168 יחידות שטח. גם מבע זה נחשב משפט צייוי, ללא התייחסות או עיגון ברעיונות קודמים של הצ'אט (היעדר עיגון). בסופו של דבר הצליח הצ'אט להגיע לתשובה הנכונה, ועל כך הגיבה נטע: "עכשיו זה נכון. הסבר איך קיבלת קודם 145?".

מעניין לשים לב שרק לאחר שהצ'אט הצליח לפתור נכון, בחרה נטע לראשונה לשאול שאלה פתוחה שמבררת את מקור הטעות. לאחר מכן הצטרפו לפסוקית זו פסוקיות מבררות אחרות, כגון "עדיין לא הבנתי למה לא ראית את הצורה כבנויה מ-5 מלבנים. למה הנחת שחסר מלבן באמצע?".

שאלות אלו בחנו במפורש את התוצאה השגויה הקודמת וחקרו את ההנחה שהוליכה אליה. נטע למעשה עברה מהנחות ניהוליות כלליות לשאלות פתוחות ומאבחנות, המבקשות להתחקות אחר תהליך החשיבה של הצ'אט. מבעים אלו קודדו לשאלה פתוחה עם מצב אפיסטמי נמוך ועיגון חלקי (Mood: open interrogative; Modality: low epistemic; Uptake: partial).

ניתוח רוחב של כלל האירועים

טבלה 2 מתארת את התפלגות סיווג הרספונסיביות של 144 הפסוקיות שנותרו לפי מטריצה דו־ממדית של פסוקיות פתוחות או סגורות ופסוקיות מעוגנות או לא מעוגנות, על פי סיווג רספונסיביות הפסוקית המפורטת בטבלה 1 לעיל.

טבלה 2: התפלגות סיווג הרספונסיביות של הפסוקיות לפי עיגון ופתיחות או סגורות

סה"כ	פסוקיות לא מעוגנות (no uptake)	פסוקיות מעוגנות (some uptake)	
			פסוקיות פתוחות
52 (~36%)	21 (~15%)	31 (~22%)	(Mood: open interrogative) OR (Mood: closed interrogative AND Modality: low)
			פסוקיות סגורות
92 (~64%)	51 (~36%)	41 (~29%)	(Mood: closed interrogative AND Modality: high) OR (Mood: declarative OR imperative)
144 (100%)	72 (~50%)	72 (~50%)	סה"כ

כפי שעולה מהטבלה, העיגון התפלג במידה שווה בין הפסוקיות: 72 קודדו לנעדרות עיגון ו-72 כבעלות עיגון. עם זה האיזון הכולל הזה לא התפלג באותה דרך בשאלת הפתיחות או סגירות. פסוקיות שקודדו לפתוחות היו כ־36% מהאירועים. מתוך פסוקיות אלו אחוז הפסוקיות המעוגנות היה גבוה יותר מהפסוקיות שלא היה בהן עיגון (22% לעומת 15%).

אם נשווה בין הפסוקיות שקודדו לפתוחות ובין הפסוקיות שקודדו לסגורות (למשל, ציווי או שאלות), הרי שב־60% (31 מתוך 52) מפסוקיות הפתוחות פרחי ההוראה עשו עיגון לחשיבה של המודל השפתי. לעומת זאת אחוז נמוך יותר של כ־45% (41 מתוך 91) מהפסוקיות הסגורות היו מעוגנות בדברי המודל. כלומר, פסוקיות פתוחות נטו להיות מעוגנות יותר מפסוקיות סגורות.

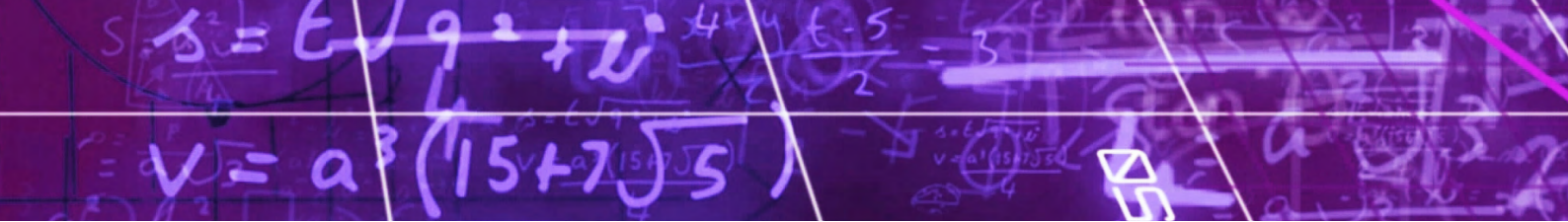
ראוי לציין כי את 31 הפסוקיות שקודדו כפתוחות ומעוגנות (ראו טבלה 2) הפיקו תשעה פרחי הוראה בלבד מתוך כלל 20 המשתתפים. ממצא זה מלמד שפיזור פסוקיות מסוג זה לא אחיד, כיוון שהן הופיעו והתקבצו בפחות ממחצית השיחות. במילים אחרות, השילוב בין פתיחות לשוניות ובין יחס ישיר כלפי ההיסק המתמטי של המודל לא שימש פרקטיקה נפוצה. הממצא המשתמע – שאר 11 פרחי ההוראה לא הציגו אף מופע של פסוקית פתוחה ומעוגנת בכל ההתכתבות עם הצ'אט – הוא ממצא עיקרי של הניתוח.

עוד עולה מטבלה 2 כי הקטגוריה של פסוקיות מעוגנות מאגדת שתי רמות: עיגון חלקי וישיר, ואחוז העיגון הישיר הוא 25% מכלל הפסוקיות (36 מתוך 144). נתון זה מלמד שרק פעמים מעטות נעשה שימוש מפורש וישיר בדברי המודל השפתי המשמש בסיס להמשך השיח, כלומר נעשה ניסיון לקחת את נקודת המבט של המודל כדי לחלצו מטעות. ממצא אחר הוא כי רק 12 מתוך 20 פרחי ההוראה הראו לפחות מופע אחד של עיגון ישיר. כלומר אף שעיגון חזק בדברי המודל אכן הופיע בקרב פרחי ההוראה, הוא לא היה דפוס רווח בכלל השיחות.

דין

במחקר זה בחנו אינטראקציות בין פרחי הוראה למודלים שפתיים מתוך התמקדות בנושא הייחודי של חילוף מטעות מתמטית. ביקשנו לבחון אילו מאפיינים רספונסיביים עולים משיחות כאשר הן מנותחות באמצעות SFL-MRT. הממצאים מעידים על שימוש נמוך בפרקטיקות רספונסיביות. רוב הפסוקיות (כ־64%) סווגו סגורות (ציווי, הצהרה או שאלה סגורה), ורק כ־36% סווגו פתוחות (שאלות פתוחות). יתרה מזו, עיגון ישיר, כלומר בנייה מפורשת של ההיסק המתמטי של מודל שפה גדול, נמצא רק ברבע מהפסוקיות ובקרב פחות ממחצית מפרחי ההוראה. הממצא שלפיו 64% מהפסוקיות נשארו סגורות על אף ההדגמה שניתנה, עשוי להעיד על אתגר מהותי של פרחי הוראה לאמץ פרקטיקות רספונסיביות לאחר התנסות קצרה, כלומר על נטייתם להשתמש בדפוסי שיח סמכותיים. נוסף על כך ייתכן שהעבודה עם מודל שפה טקסטואלי, בניגוד לדינמיקה האנושית עם תלמיד, מעודדת בשלב הראשון דפוסי שיח סגורים יותר.

חשוב לציין שפסוקיות פתוחות בפני עצמן לא הבטיחו אינטראקציה רספונסיבית. כמה מהשאלות הפתוחות היו כלליות ופרוצדורליות ולא נשענו על ההיסק המתמטי הייחודי שהציג המודל השפתי. הבדל זה עשוי לשקף גישות מגוונות של פרחי ההוראה כלפי מודל השפה עצמו. למשל, כמה מפרחי ההוראה ראו במודל שותף לשיח וחשבו כי החשיבה המתמטית שלו מצדיקה פרשנות ותגובה. לעומתם אחרים ראו במודל כלי טכני שמתפקד היטב כאשר הוא מונחה באמצעות ציווי.



חקר המקרה שהצגנו מציע גישה שלפיה רספונסיביות אינה מאפיין קבוע הטמון ביחיד אלא תהליך דינמי שעשוי להתפתח תוך כדי שיחה. בדוגמה שהצגנו התקשורת של פרח ההוראה, נטע, התאפיינה תחילה בציוויים סגורים נעדרי עיגון שבהם ניסתה להכתיב למודל את דרך הפתרון במקום לאתר ולהיכנות על ההיסק השגוי שלו. מעניין לשים לב כי רק לאחר שהצ'אט הגיע לפתרון הנכון, נטע הציגה הוראה רספונסיבית שהתאפיינה בשאלת שאלות פתוחות, ממוקדות אבחון, מתוך התייחסות ישירה לטעות שעשה המודל. מעבר זה מעיד שיש ביכולתה לאמץ את נקודת המבט של האחר או להיות סקרנית כלפיה, אך היא לא השתמשה בפרקטיקה זו כדי לחלץ את המודל מטעות, אלא רק כדי להבין את המקור לטעות, לאחר החילוך. בפרשנות הממצאים יש לזכור שפרחי ההוראה לא קיבלו הנחיה מפורשת להיות רספונסיביים. אנו משערים שמודעות מכוונת לממדים של הוראה רספונסיבית עשויה הייתה לשנות את אופי האינטראקציות ולהגיע לתוצאות אחרות מאלו שהתקבלו במחקר הגישוש הזה.

מבחינה מתודולוגית, מחקר זה תורם לפיתוח ה-MRT-SFL, מסגרת שתשמש לניתוח רספונסיביות מתמטית באינטראקציות בין מורים לכלים מבוססי מודלי שפה גדולים. השימוש במסגרת תאורטית לשונית, במקרה שלנו – SFL, מאפשר להגדיר הגדרה אופרטיבית היבטים מגוונים של תקשורת מתמטית, ומתוך כך לראות ברספונסיביות לא תכונה כללית, או מקושרת למהלכים ספציפיים (כמו אצל Jacob & Empson, 2016) אלא תהליך דיאלוגי.

מגבלה מרכזית של המחקר, שהיא גם מאפיין מהותי שלו, היא שפרחי ההוראה תקשרו עם מודלי שפה כלליים (ושונים זה מזה), ולא עם כלי פדגוגי ייעודי, כגון סימולטור שתוכנן במיוחד לדמות חשיבה ותגובות של תלמידים. עיצוב ייעודי של הפעילות עשוי להשפיע על אופי הדיאלוגים. עם זאת, נקודה זו משקפת גם את הרציונל של המחקר: בטרם משקיעים משאבים מהותיים בפיתוח סימולציות פדגוגיות ייעודיות, יש חשיבות לבחון את האינטראקציות כפי שהן מתנהלות בטבעיות בין מורים למערכות כלליות של מודל שפה זמין. מתוך כך המחקר מספק נקודת מוצא (base-line) אמפירית ראשונית, היכולה לשמש תשתית לפיתוח מושכל יותר של כלים פדגוגיים עתידיים המותאמים למטרות של הכשרת מורים למתמטיקה.

הממצאים מלמדים על שלושה כיווני מחקר עתידיים. ראשית, מחקר התערבות שיבחן האם הנחיה מפורשת לנהל דיאלוג רספונסיבי, וחשיפה מוקדמת להבדלים מסוימים בממדי ה-Mood והעיגון לפני האינטראקציה עם מודל שפה או במהלכה, משנים את אופי הדיאלוג לכיוון רספונסיבי יותר. שנית, פיתוח וחקר של סימולטורים פדגוגיים המדמים חשיבה של תלמידים ומציגים טעויות נפוצות, שיאפשרו בחינה ממוקדת יותר של פרקטיקות חילוך מטעות (Schwartz & Czaczkes, 2026). שלישית, מחקר השוואתי שיבחן האם וכיצד רספונסיביות המתפתחת באינטראקציות עם מערכות מודל שפה גדול עוברת גם לאינטראקציות עם תלמידים אמיתיים בכיתה. שאלה זו מצויה בלב הפוטנציאל הפדגוגי של שימוש במודלי שפה גדולים לצורכי הכשרת מורים, שכן היא נוגעת ליכולת של התנסויות מבוססות מודל שפה לשפר פרקטיקות הוראה בפועל.

לסיכום, זהו מחקר ראשוני הבוחן את דרכי השיח של פרחי הוראה באינטראקציה עם מודלי שפה בנושא טעויות מתמטיות, ומציע תשתית אמפירית ורעיונית להמשיך לפתח ולהעמיק שימושים פדגוגיים-מתמטיים בסביבות מבוססות בינה מלאכותית. אנו מקווים שהממצאים והתובנות המוצגים כאן ישמשו בסיס לפיתוח מחקרים עתידיים שיתרמו לעיצוב התערבויות, כלים וסביבות למידה יעילות לצורך פיתוח פרקטיקות של הוראה רספונסיבית במתמטיקה.

רשימת מקורות

- Bishop, J. P., Hardison, H. L., & Przybyla-Kuchek, J. (2022). Responsiveness to students' mathematical thinking in middle-grades classrooms. *Journal for Research in Mathematics Education*, 53(1), 10–40. <https://doi.org/10.5951/jresmetheduc-2020-0188>
- Borasi, R. (1994). Capitalizing on errors as “springboards for inquiry”: A teaching experiment. *Journal for Research in Mathematics Education*, 25(2), 166–208. <https://doi.org/10.5951/jresmetheduc.25.2.0166>
- Bray, W. S. (2011). A collective case study of the influence of teachers' beliefs and knowledge on error-handling practices during class discussion of mathematics. *Journal for Research in Mathematics Education*, 42(1), 2–38. <https://doi.org/10.5951/jresmetheduc.42.1.0002>
- Brodie, K. (2014). Learning about learner errors in professional learning communities. *Educational Studies in Mathematics*, 85(2), 221–239. <https://doi.org/10.1007/s10649-013-9507-1>
- Buber, M. (1958). *I and Thou* (2nd ed., R. G. Smith, Trans.). Scribner. (Original work published 1923)



- Cassinadri, G. (2024). ChatGPT and the technology-education tension: Applying contextual virtue epistemology to a cognitive artifact. *Philosophy & Technology*, 37(1), Article 14. <https://doi.org/10.1007/s13347-024-00701-7>
- Ebbelind, A. (2023). A functional view on language: A methodology for mathematics education to study shifts in prospective teachers' discursive patterns. *International Journal of Mathematical Education in Science and Technology*, 54(8), 1731–1745. <https://doi.org/10.1080/0020739X.2023.2204506>
- Ergene, Ö., & Çaylan-Ergene, B. (2026). Pre-service mathematics teachers' and engineering students' perceptions of ChatGPT in mathematics: Development, validation and implementation study. *Digital Experiences in Mathematics Education*, 12, 68–94. <https://doi.org/10.1007/s40751-025-00176-x>
- Halliday, M., & Matthiessen, C. M. (2004). *Introduction to functional grammar* (3rd ed.). Hodder Arnold.
- Jacobs, V. R., Lamb, L. L., & Philipp, R. A. (2010). Professional noticing of children's mathematical thinking. *Journal for Research in Mathematics Education*, 41(2), 169–202. <https://doi.org/10.5951/jresmetheduc.41.2.0169>
- Jacobs, V. R., & Empson, S. B. (2016). Responding to children's mathematical thinking in the moment: an emerging framework of teaching moves. *ZDM – Mathematics Education*, 48(1–2), 185–197. <https://doi.org/10.1007/s11858-015-0717-0>
- Lee, S., Lee, N., & Sah, Y. J. (2020). Perceiving a mind in a chatbot: Effect of mind perception and social cues on co-presence, closeness, and intention to use. *International Journal of Human-Computer Interaction*, 36(10), 930–940. <https://doi.org/10.1080/10447318.2019.1699748>
- Lin, Y.-F., Yang, E. F.-Y., Wu, J.-S., Yeh, C. Y. C., Liao, C.-Y., & Chan, T.-W. (2025). Enhancing students' authentic mathematical problem-solving via error analysis learning activity using GPT-4. *Research and Practice in Technology Enhanced Learning*, 20, Article 34. <https://doi.org/10.58459/rptel.2025.20034>
- Marco, N., & Stylianides, A. J. (2025). ChatGPT's mathematical knowledge and the nature of human–AI collaboration in mathematical inquiry. *International Journal of Mathematical Education in Science and Technology*, 56(11), 2279–2297. <https://doi.org/10.1080/0020739X.2025.2543817>
- Mejía-Ramos, J. P., Fuller, E., Weber, K., Rhoads, K., & Samkoff, A. (2012). An assessment model for proof comprehension in undergraduate mathematics. *Educational Studies in Mathematics*, 79(1), 3–18. <https://doi.org/10.1007/s10649-011-9349-7>
- Rizos, I., & Gkrekas, N. (2025). The impact of LLMs on mathematics education and research at the university. *Social Sciences & Humanities Open*, 12, Article 101969. <https://doi.org/10.1016/j.ssaho.2025.101969>
- Robertson, A. D., Scherr, R. E., & Hammer, D. (Eds.). (2016). *Responsive teaching in science and mathematics*. Taylor & Francis. <https://doi.org/10.4324/9781315689302>
- Schwartz, G. (2025). Moving beyond moves: Redefining mathematically responsive teaching as a humanistic stance. In C. Cornejo, P. Felmer, D. M. Gómez, P. Dartnell, P. Araya, A. Peri, & V. Randolph (Eds.), *Proceedings of the 48th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 267–274). PME.
- Schwartz, G., & Czaczkes, S. (2026, June 3–5). *Interactional formative feedback in AI-based simulations for assessing mathematically responsive teaching* [Conference presentation]. FAME2 – The Second Feedback and Assessment in Mathematics Education conference, Budapest, Hungary.
- Selden, A., & Selden, J. (2003). Validations of proofs considered as texts: Can undergraduates tell whether an argument proves a theorem? *Journal for Research in Mathematics Education*, 34(1), 4–36. <https://doi.org/10.2307/30034698>
- Silver, E. A. (1997). Fostering creativity through instruction rich in mathematical problem solving and

problem posing. *ZDM – Mathematics Education*, 29(3), 75–80. <https://doi.org/10.1007/s11858-997-0003-x>

- Stylianides, A. J., & Stylianides, G. J. (2024). Important mathematical competences for students in the age of generative artificial intelligence: Proof comprehension and evaluation, and ChatGPT. *Ευκλείδης γ'* [Euclid C], 100(1–2). https://hms.gr/wp-content/uploads/2025/04/4_0_29_INPRESS_StylianidesStylianides_EN_14.pdf
- Taani, O., & Alabidi, S. (2024). ChatGPT in education: Benefits and challenges of ChatGPT for mathematics and science teaching practices. *International Journal of Mathematical Education in Science and Technology*, 56(9), 1748–1777. <https://doi.org/10.1080/0020739X.2024.2357341>
- Wang, R. E., Zhang, Q., Robinson, C., Loeb, S., & Demsky, D. (2024). Step-by-step remediation of students' mathematical mistakes. *arXiv:2310.10648*. <https://doi.org/10.48550/arXiv.2310.10648>
- Zhuang, Y., & Zhang, S. (2025). Integrating ChatGPT in mathematics teacher education: AI-based simulation role-playing to support practice-based teaching. *International Journal of Artificial Intelligence in Education*, 35(6), 3873–3895. <https://doi.org/10.1007/s40593-025-00519-0>