

# מהאנושי למכונה

אתגרים והזדמנויות בהערכת הבעה בכתב  
באמצעות בינה מלאכותית

ד"ר תמי סבג שושן, הדיל שחברי, ד"ר הדס גלברט, בועז רוזנבאום  
הרשות הארצית למדידה והערכה (ראמ"ה)





## חלק 3

השפעת הבינה  
המלאכותית



## חלק 2

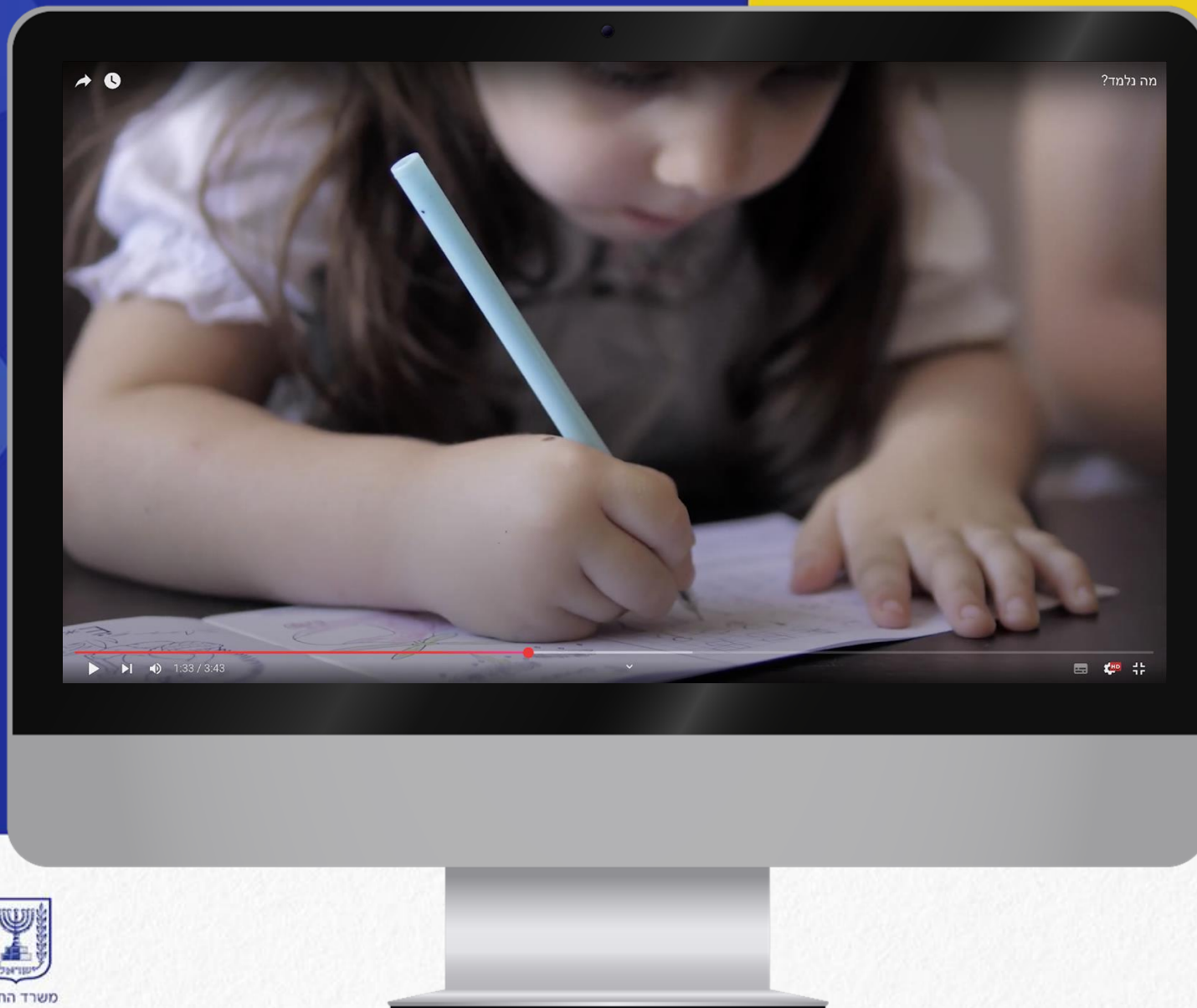
הערכה מבוססת  
בינה מלאכותית



## חלק 1

האתגרים  
בהערכת כתיבה

# חשיבות הכתיבה



# מורכבות הערכת הכתיבה

מצריך  
זמן רב

מצריך ידע  
מקצועי  
נרחב

נתון  
לפרשנות  
בהערכה

הבעיה  
תלמידים רבים מתקשים בכתיבה



המענה  
יוצרים הזדמנויות לכתיבה  
ולשיחה על טקסטים



משימה לדוגמה

# הבעה בכתב לכיתה ח' תיאור המשימה



בכתבה שפורסמה בעיתון בית הספר צוטטו תלונות הורים בנוגע לתקשורת בעידן הדיגיטלי של ימינו.

לדברי ההורים, בני הנוער יוצרים קשר עם חבריהם בעיקר באמצעים דיגיטליים, כגון רשתות חברתיות, טלפון נייד או משחקי מחשב שיתופיים ברשת. הם מבלים זמן רב מדי מול מסכים וממעטים להיפגש עם חבריהם פנים אל פנים.

1. כתוב את עמדתך לקוראי עיתון בית הספר בנוגע לשאלה איזו תקשורת עדיפה ומדוע היא עדיפה.
2. נמק את עמדתך והבא דוגמאות התומכות בה.
3. כתוב כ 150-מילים.

לשון

מבנה

תוכן

” את צריכה ללמד  
איש חכם מאוד  
שלא יודע דבר בתחום..”

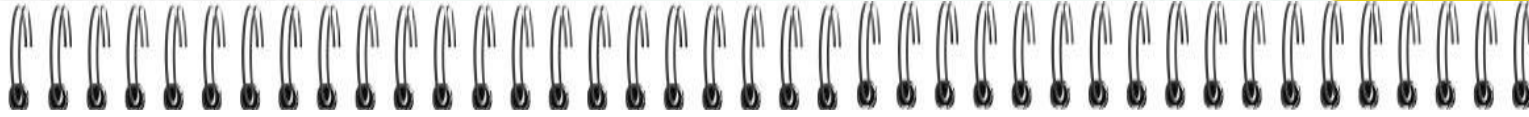
המבדקים שבודקים  
את עצמם!

אימון הבינה  
המלאכותית  
להערכה



# התאמת המחוון המשגה > שפה מקצועית

מה שברור למומחה צריך ללמד את המכונה



## 2 התוכן ומבנה הטקסט

- |   |  |  |   |
|---|--|--|---|
| <input type="checkbox"/> התוכן ומבנה הטקסט אינם הולמים את ה | <input type="checkbox"/> התוכן ומבנה הטקסט הולמים את דרישות המשימה במידה חלקית | <input type="checkbox"/> התוכן ומבנה הטקסט הולמים את דרישות המשימה במידה רבה | <input type="checkbox"/> התוכן ומבנה הטקסט הולמים את דרישות המשימה במידה רבה מאוד |
|---|--|--|---|

---

---

---

---

---

---

---

---

---

---

הציון (15-0 נק):

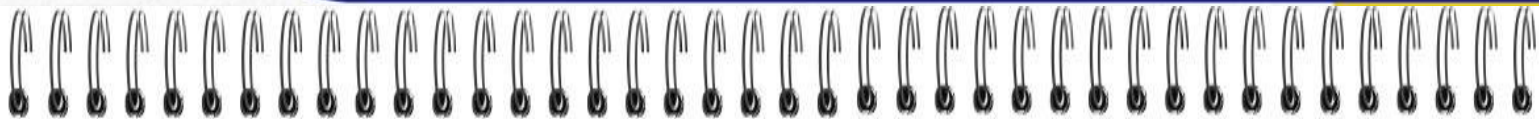
- האם יש בטקסט פתיחה וסיום? באיזו מידה הם איכותיים? (2-0 נק')
- האם יש בטקסט עמדה מבוססת ומוצדקת? באיזו מידה רכיבי התוכן (עמדה, ביסוס, הצדקה) איסתיים? (7-0 נק')
- באיזו מידה הטקסט לכיד ומקושר? (4-0 נק')
- באיזו מידה הטקסט עצמאי וניכר שיש בו חזרות לנמען ולמקום הפרסום? (2-0 נק')

**מה משמעות עמדה מבוססת ומנומקת?**

**מה משמעות טקסט עצמאי? לכיד? מקושר?**



# סוגיות הקשורות למחון > שפה ותרבות



משלב הלשון  
 ואוצר המילים  
 אינם הולמים את  
 דרישות המשימה

---



---



---

משלב הלשון  
 הולמים את  
 דרישות המשימה  
 במידה חלקית

משלב הלשון  
 הולמים את  
 דרישות המשימה

**משלב הלשון  
 ואוצר המילים**  
 (0-3 נק')

– האם משלב הלשון  
 ביוני-גבוה בכל הטקסט?  
 – באיזו מידה אוצר המילים  
 בטקסט עשיר, מגוון ומדויק?

**משלב לשון של  
 תלמידי כיתה  
 ה' / ח'?**

יש שגיאות כתיב  
 רבות, והן פוגמות  
 בהבנת הטקסט

---



---



---

יש שגיאות  
 כתיב רבות

יש מעט  
 שגיאות כתיב

אין שגיאות  
 כתיב

**הכתיב**  
 (0-2 נק')

– באיזו מידה הטקסט כתוב  
 ללא שגיאות כתיב?  
 – באיזו מידה שגיאות הכתיב  
 פוגמות בהבנת הטקסט?

**מה נחשבת  
 כשגיאה?  
 מה בין שגיאת  
 כתיב-לבין שגיאת  
 דקדוק**

הציון (0-2 נק'):



## 2 התוכן ומבנה הטקסט

- התוכן ומבנה הטקסט אינם הולמים את המשימה
- התוכן ומבנה הטקסט הולמים את דרישות המשימה במידה חלקית
- התוכן ומבנה הטקסט הולמים את דרישות המשימה במידה רבה
- התוכן ומבנה הטקסט הולמים את דרישות המשימה מאוד

איכות התוכן ואיכות מבנה הטקסט (15-0 נק')

- האם יש בטקסט פתיחה וסיום? באיזו מידה הם איכותיים? (2-0 נק')
- האם יש בטקסט עמדה מבוססת ומוצדקת? באיזו מידה רכיבי התוכן (עמדה, ביסוס, הצדקה) איסתיים? (7-0 נק')
- באיזו מידה הטקסט לכיד ומקושר? (4-0 נק')
- באיזו מידה הטקסט עצמאי וניכר שיש בו מודעות לנמען ולמקום הפרסום? (2-0 נק')

מה הופך רכיב לאיכותי?

מה קורה אם יש רק פתיחה ואין סיכום? או שיש פתיחה איכותית אבל אין סיכום?

כיצד מציינים בהינתן טווח ציונים? כמה נקודות לעמדה? כמה לנימוקים?

הציון (15-0 נק):

## אימון הבינה

## שלבי אימון הבינה:

- ✓ הוראות
- ✓ הוראות מפורטות
- ✓ דוגמאות מנותחות
- ✓ הוראות דירוג
- ✓ טיוטת מחשבות
- ✓ הנחיות לכתיבת המשוב



# תהליך דיוק המחוננים



## הגדרה:

משפט אחד או מספר משפטים המופיעים בתחילתו של הטקסט העוזרים לקורא להבין במה הטקסט עוסק.

## הוראות כלליות:

- האם ישנה בטקסט פתיחה המהווה רקע או הקדמה לטיעון?
- האם הפתיחה מרחיבה ומעשירה את הטקסט מעבר לתפקידה?

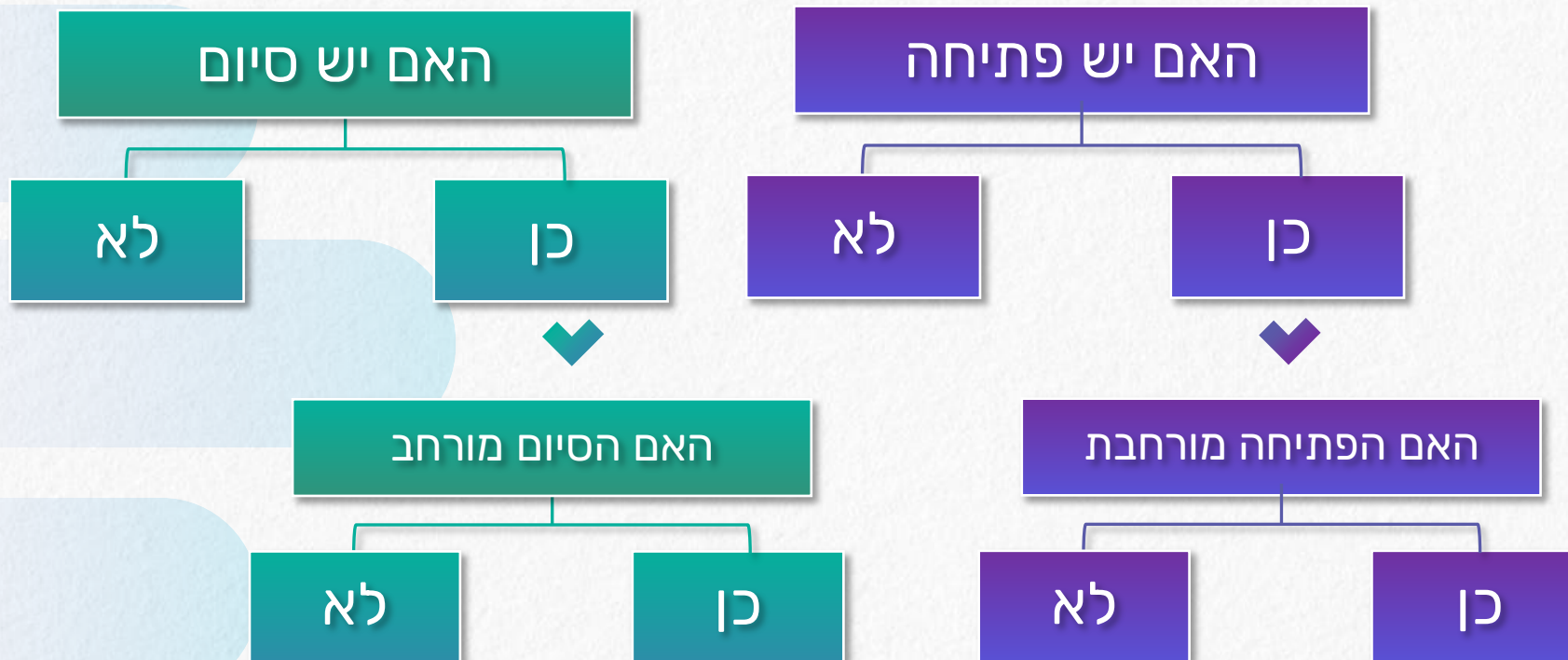


המחווון  
המותאם  
בינה

– האם יש בטקסט פתיחה  
וסיוס? באיזו מידה הם  
איכותיים?  
(2-0 נק')



המחווון  
המקורי



# רכיבי הערכה פתיחה וסיום



אם חסרים בטקסט גם פתיחה וגם סיום	אם ישנה רק פתיחה או רק סיום	אם קיימים גם פתיחה וגם סיום ושניהם ברמת בסיס	אם קיימים גם פתיחה וגם סיום ולפחות אחד מהם ברמה מתקדמת
	0 נקודות		

כיום יש מגוון רחב של אפשרויות לתקשר אחד עם השני כגון: רשתות חברתיות, טלפון נייד, משחקי מחשב שיתופיים ועוד ...  
בגלל דרכים אלו הילדים מבליים הרבה זמן מול המסך ומקדישים מעט מזמנם להיפגש פנים מול פנים עם חברים. בשל כך הורים רבים מתלוננים לגבי זה שהילדים שלהם מבליים את רוב זמנם מול מסכים ומכשירים דיגיטליים .....

**הפתיחה מציגה את הנושא בצורה מעמיקה ומהווה רקע טוב לטיעון. הסיום מסכם את עמדת הכותב וכולל המלצה לפעולה ומסקנה, ובכך מרחיב מעבר לתפקידו הבסיסי.**

בשנים האחרונות, ההורים מתלוננים על בני נוער שמבליים את רוב הזמן במסך וממעטים להיפגש עם חבריהם פנים על פנים. ....

**הטקסט כולל פתיחה וסיום ברמת בסיס. הפתיחה מציגה את הנושא והבעיה, והסיום מסכם את עמדת הכותבת בנושא.**



העמקת  
המשגה  
ודיוק



המחווה  
והמשימה  
נפתחו  
לדיון  
מחודש

ציינון ברור  
ואחיד

✓ קיים >  
מורחב  
✓ נמוך > בסיס  
> מתקדם



ארגון  
הרכיבים  
מחדש







נימוק: הטקסט כתוב כמכתב אישי מהטוס לחבריו, מתייחס לקשיים ולטיפול אצל דוליטל ומפרט אותם היטב. כמו כן, מופיעה המלצה מנומקת ומפורטת להיבדק אצל דוליטל. לכן הטקסט ראוי לציון המקסימלי.

איכות התוכן  
- באיזה מידה הטקסט הכתוב מותאם למשימה המתבקשת?  
- באיזו מידה שלושת הרכיבים המתבקשים כתובים...  
7 / 7

נימוק: הטקסט עוסק כולו בנושא המרכזי תוך שימוש מנוון ומדויק במילות קישור ובמאזכרים. המשפטים מאורגנים בצורה הנייט ויש זרימה טובה וקשר ברור בין הרעיונות.

לכידות וקישוריות  
- באיזו מידה הטקסט לכיד?  
- באיזו מידה הטקסט מקושר?  
4 / 4

נימוק: מבנה המשפטים בטקסט תקין והמילים תואמות במין (זכר/נקבה), במספר (יחיד/רבים), בזמן (עבר/הווה/עתיד), וביידוע (הוספת ה' הידיעה).

תחביר  
- באיזו מידה מבנה המשפטים תקין?  
- באיזו מידה יש התאם דקדוקי במין, מספר, וביידוע?  
2 / 2

נימוק: בכתיבתך השתמשת בסימני הפיסוק הבאים: לוסן, פסיק, נקודה, סימן קריאה, סימן שאלה, גרשיים, מקף, נקודותיים. ניכר שימוש מספק בסימני פיסוק לאורך הטקסט.

סימני פיסוק  
- האם יש בטקסט שימוש נכון ומתאים לתפקידיהם של סימני הפיסוק? (נקודה, סימן שאלה, נקודתיים, מירכאות, סימן קריאה)  
2 / 2

נימוק: אוצר המילים במכתב עשיר ומנוון, עם שימוש במילים מדויקות ומתאימות לתוכן. המשלב הלשוני נבנה לכל אורך הטקסט ומתאים מאוד לטקסט כתוב.

אוצר מילים  
- באיזו מידה אוצר המילים בטקסט עשיר, מנוון ומדויק?  
- באיזו מידה משלב הלשון מתאים לטקסט כתוב ולתוכן הטקסט?  
2 / 2

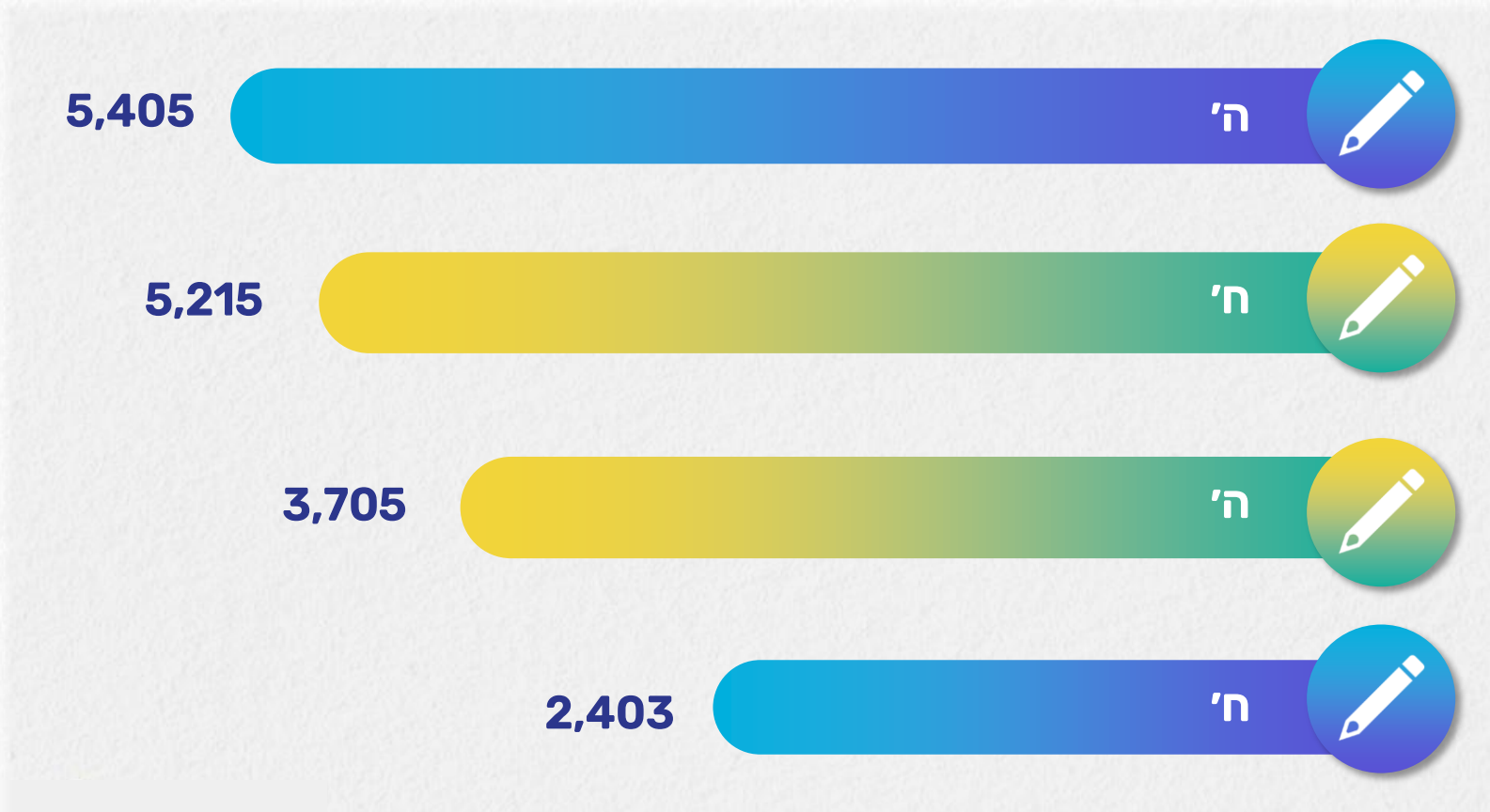
נימוק: הכתיב תקין

כתיב  
- באיזו מידה הטקסט כתוב באופן תקין ללא שגיאות כתיב?  
- באיזה מידה שגיאות הכתיב פוגמות בהבנת הטקסט?  
2 / 2

# מסך משוב

תפקיד הבינה  
ותפקיד המורה

# מספר משיבים למבחני הבעה בכתב תשפ"ד



ערבית



עברית



## נתונים מתוך ניסיון שימוש בכלי הדיגיטלי

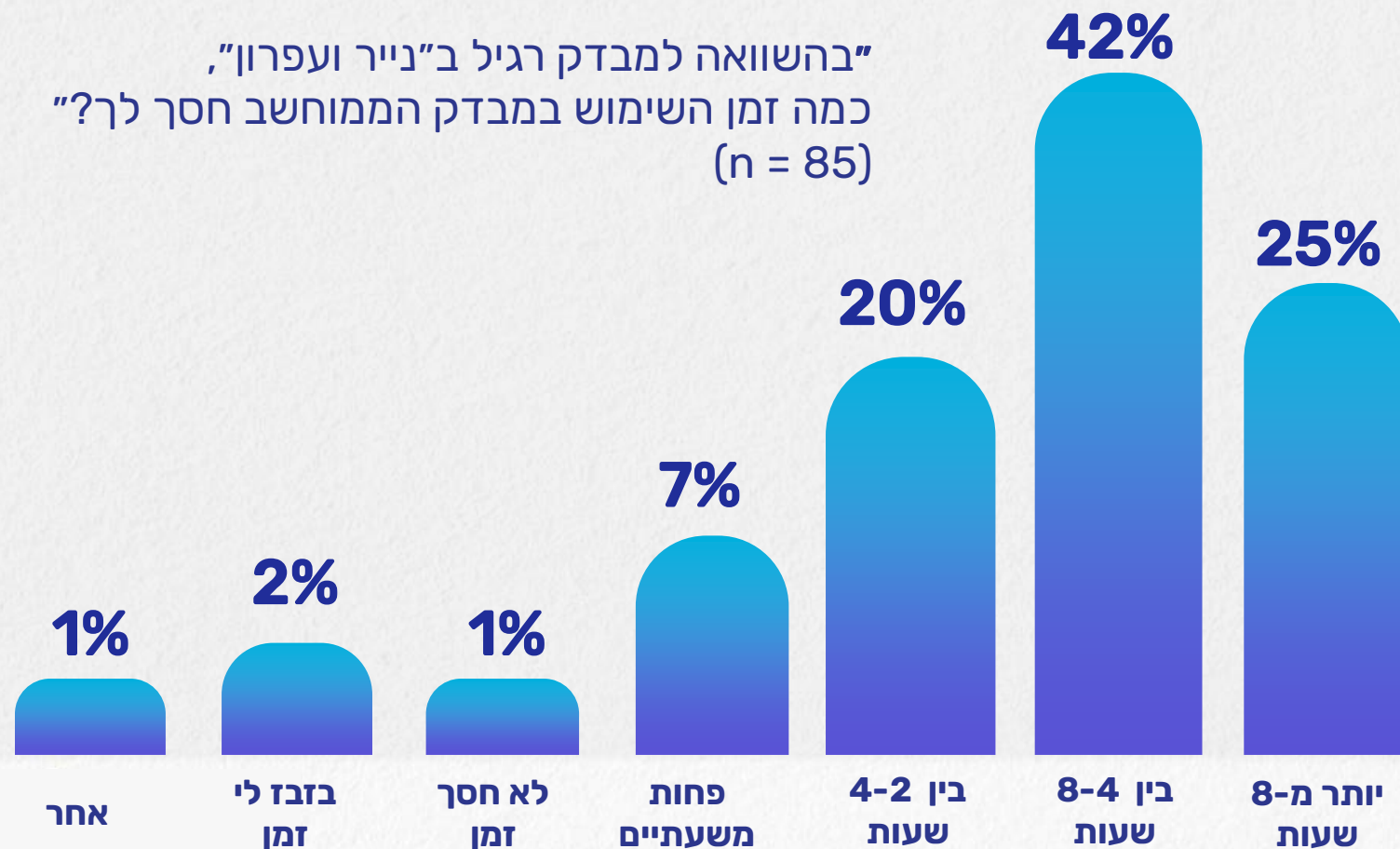


שני שלישי מדווחים שנחסך להם בין חצי יום ליום עבודה (ואף יותר מכך) מעבודתם

## אתגר הזמן

נתונים  
מתוך ניסיון  
שימוש בכלי  
הדיגיטלי

"בהשוואה למבדק רגיל ב"נייר ועפרון",  
כמה זמן השימוש במבדק הממוחשב חסך לך?"  
(n = 85)



סך מקבלי הסטורי (N = 1,198)

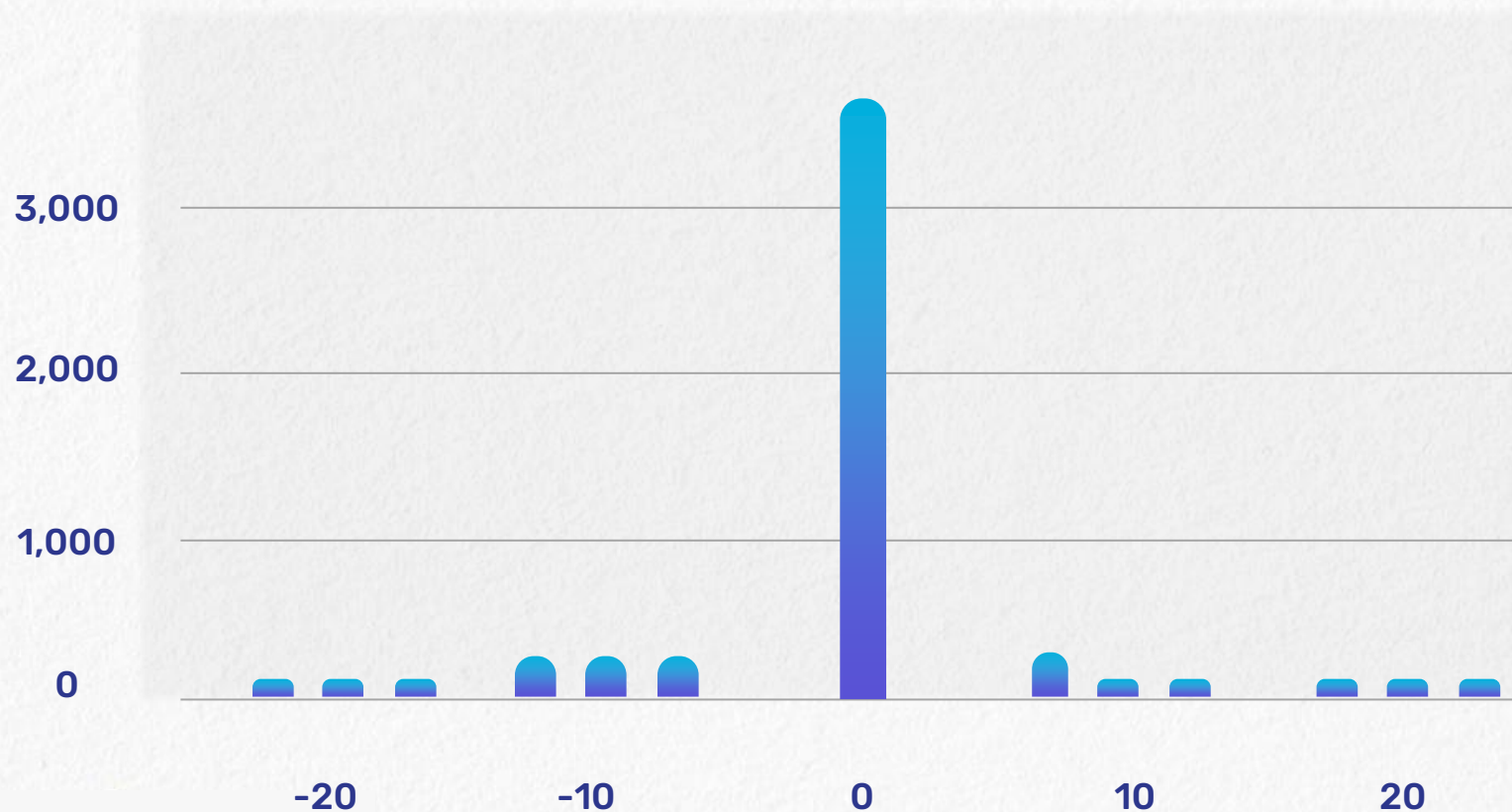


# אתגר הידע

נתונים  
מתוך ניסיון  
שימוש בכלי  
הדיגיטלי



מרבית המורים לא עורכים את הציונים של הבינה

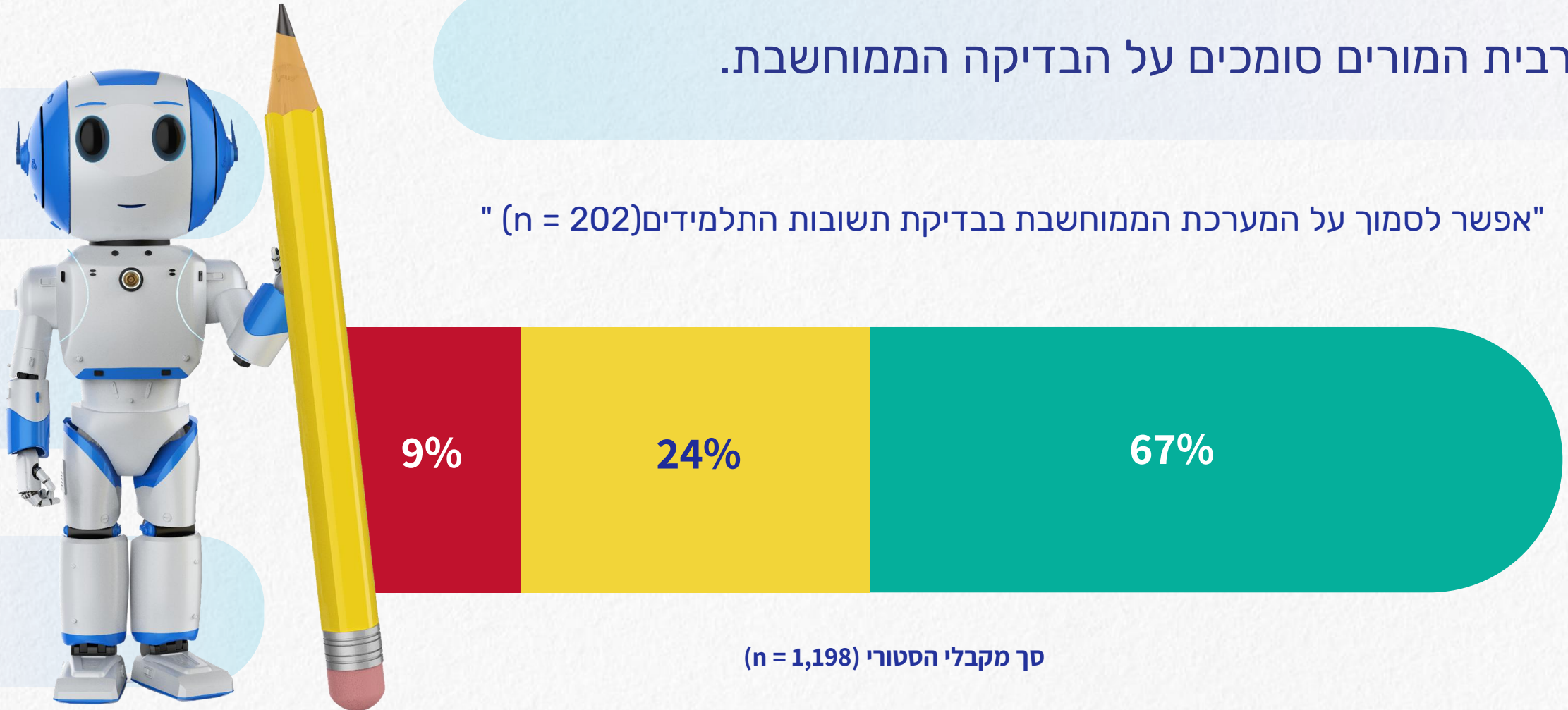


נקודות שנוספו או הורדו (0-100)

# נתונים מתוך ניסיון שימוש בכלי הדיגיטלי

מרבית המורים סומכים על הבדיקה הממוחשבת.

"אפשר לסמוך על המערכת הממוחשבת בבדיקת תשובות התלמידים (n = 202)"



סך מקבלי הסטורי (n = 1,198)

# מסקנות מתוך מחקר תוקף

## חישוב מדד קאפה הממושקל

0.81-1.00	0.61-0.80	0.41-0.60	0.21-0.40	0.00-0.20
הסכמה רבה	הסכמה טובה	הסכמה מתונה	הסכמה מעטה	הסכמה נמוכה

0 - מייצג הסכמה אקראית לחלוטין  
1 - מייצג הסכמה מושלמת בין המעריכים

## שיטת הבדיקה

נדגמו כ - 100 תגובות  
תלמידים מכל דרגת כיתה.

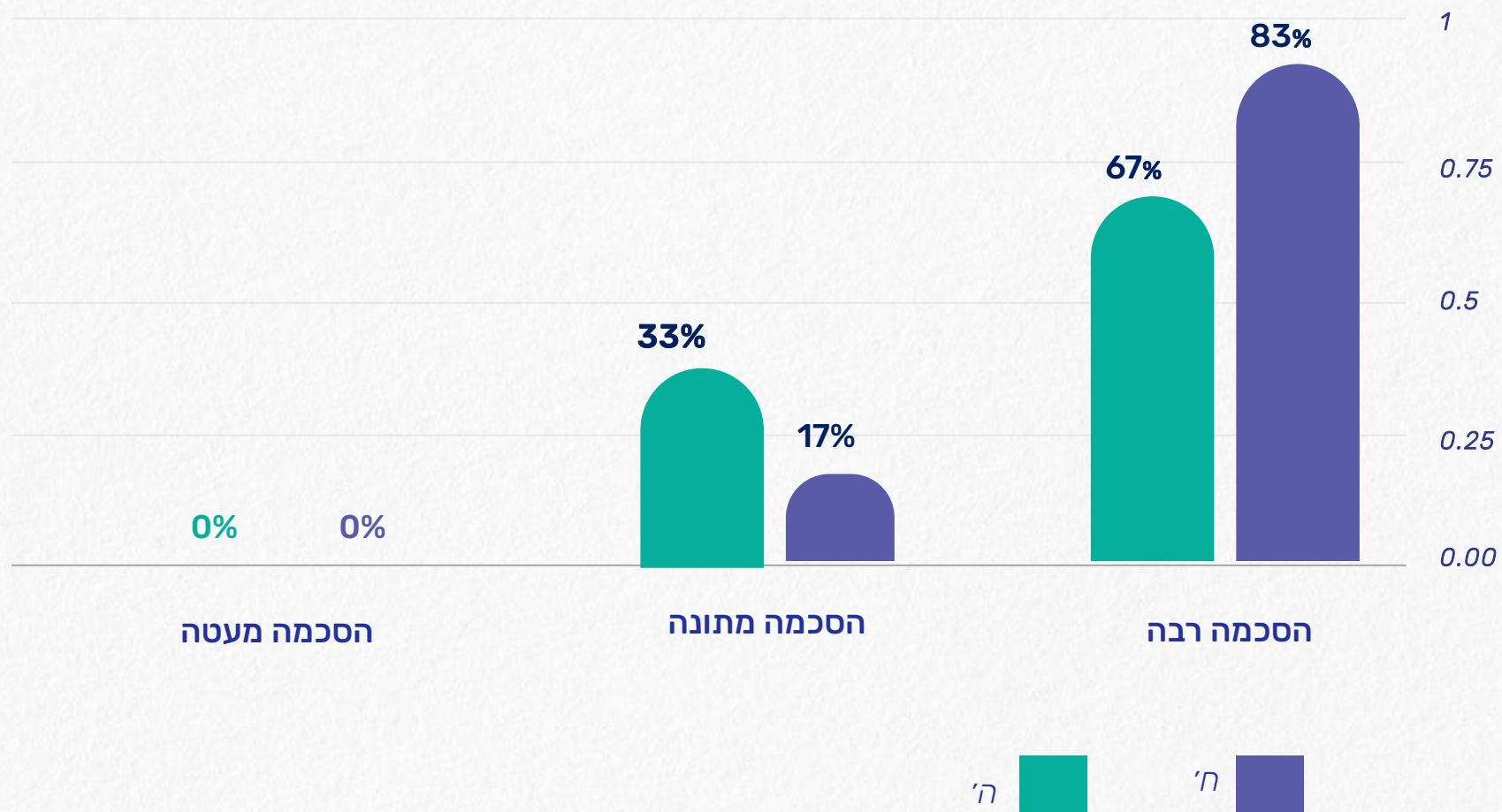
כל תגובה נבדקה במדדי  
הערכה השונים באופן  
עצמאי על ידי שני  
מעריכים, באופן בלתי תלוי.

חושבה מהימנות  
- בין מעריך א' למעריך ב'  
- בין המעריכים לבינה  
מלאכותית

## מטרת המחקר

המחקר בודק את מידת  
ההתאמה בין בדיקות  
של בינה מלאכותית  
לבין בדיקות של  
מעריכים אנושיים  
במבדקי הבעה בכתב.

# מתוך ממצאי תוקף: מידת ההסכמה בין המעריכים לבינה טובה- גבוהה



# מתוך ממצאי תוקף: הבינה מדויקת יותר בהערכת תוכן בהשוואה להערכת דקדוק ומבנה

הסכמה בין מעריכים ובין הבינה -  
הבעה בכתב שפת אם עברית כיתה ה'

מעריכים ובינה	מעריכים	
רבה	רבה	איכות התוכן
מתונה	רבה	לכידות
טובה	רבה	סימני פיסוק
טובה	רבה	אוצר מילים
מתונה	רבה	כתיב
מתונה	רבה	דקדוק

הסכמה בין מעריכים ובין הבינה -  
הבעה בכתב שפת אם עברית כיתה ח'

מעריכים ובינה	מעריכים	
ח'	ח'	
טובה	רבה	פתיחה וסיום
טובה	רבה	גוף הטקסט
טובה	רבה	לכידות
מתונה	רבה	טקסט מותאם לסוגה
טובה	רבה	אוצר מילים
טובה	רבה	כתיב
מתונה	רבה	דקדוק

לעיון מעמיק בדו"ח < אתר רמ"ה



## אתגרים שבוצעו

נתון  
לפרשנות  
בהערכה



מצריך  
זמן רב



מצריך ידע  
מקצועי  
נרחב,  
מיומנות



## אתגרים חדשים

התייחסות  
למסרים  
רגשיים

השלכות  
פדגוגיות

מענה  
לאתגרים  
קיימים  
יוצר אתגרים  
חדשים





כדי להבטיח שילוב מוצלח  
של AI בהערכת הבעה בכתב,  
יש לאמץ גישה מאוזנת,

שמשכללת את תפקיד המורה  
ומאפשרת לו להתמקד בפן האנושי  
והפדגוגי תוך ניצול היתרונות  
הטכנולוגיים של המערכת.





# תודה רבה!

