

מבחן "תנופה למחר"

מסגרת מושגית ודוח טכני על תהליך הפיתוח

ד"ר תמי סבג-שושן ואימאן עואדיה (אגף פיתוח כלים ברמ"ה)

בשנת 2020, לאחר תהליך ארוך ומעמיק, גיבש משרד החינוך בהובלת המזכירות הפדגוגית ואגף אסטרטגיה ותכנון את "[תפיסת הלמידה 2031-2021](#)". במסגרתה, הוגדרו חמישה תחומי ידע ליבתיים, 13 מיומנויות ויכולות יסוד וחמישה ערכים שמערכת החינוך מבקשת להנחיל לבוגריה. קידום הידע, המיומנויות והערכים שהוגדרו במסמך הוא מדד תוצאה מרכזי של משרד החינוך. רמ"ה, כמי שאמונה על מדידת תוצאות המשרד, פועלת לפיתוח כלי הערכה שיאפשרו זאת. כיום אנו בודקים באופן סדיר את כל חמשת תחומי הידע, שבע מ-13 המיומנויות ו-80% מחמשת הערכים.

מבחן "תנופה למחר" נועד לבדוק את ארבע [המיומנויות הקוגניטיביות חוצות התחומים](#) שהוגדרו במפורט בתפיסה, לאחר עבודת מטה שהשתתפו בה עשרות רבות של מנהלים במשרד: אוריינות מידע, אוריינות דיגיטלית, חשיבה יצירתית וחשיבה ביקורתית.

בדוח הטכני לסיכום שלב א' של פיתוח המבחן מתוארים המסגרת המושגית שלו, תהליך הפיתוח והתוצאות המרשימות של המחקר החלוץ. לנוכח העובדה שמדובר בפיתוח חלוצי ופורץ דרך, גם ברמה הבין-לאומית, סברנו שנכון לשקף את התהליך שנעשה ובכך לסייע לכל העוסקים בתחום.

ראשי פרקים

2	תקציר
4	תהליך העבודה
7	יסודות המבחן
10	המסגרת המושגית
17	תהליך הפיתוח
23	תוצאות המחקר החלוץ
26	מקורות והפניות
28	אנשי הצוות

תקציר

הצורך במבחן

פיתוח מבחן "תנופה למחר" הוא חלק מיוזמה מערכתית של משרד החינוך, המבוססת על מסמך "דמות הבוגר" ותפיסת הלמידה המתחדשת. הוא מושתת על ההבנה העמוקה כי מערכת החינוך וכלי ההערכה חייבים להתאים את עצמם למציאות המשתנה של המאה ה-21. בעולם שבו הידע זמין ונגיש בקלות, בעיקר בעידן המהפכני של הבינה המלאכותית היוצרת, המיקוד עובר משינון מידע לפיתוח מיומנויות חשיבה כלליות חוצות תחומים ורלוונטיות למגוון הקשרים.

המבחן נועד למדוד מיומנויות קוגניטיביות חיוניות: חשיבה ביקורתית, חשיבה יצירתית, אוריינות מידע ואוריינות דיגיטלית. הוא משמש מענה לצורך הדחוף של מערכת החינוך להכשיר את התלמידים להתמודד עם אתגרי העתיד ועם המציאות הטכנולוגית העכשווית, הכוללת את ההשפעה המהפכנית של הבינה המלאכותית היוצרת. כלי הערכה זה נועד לספק תמונה עדכנית ומדויקת של יכולות התלמידים, כדי לאפשר למשרד החינוך לכוון את הפדגוגיה ולהכין את הדור הבא לעולם שבו היכולת להשתמש בכלים ובמשאבים כדי לפתור בעיות מורכבות היא מיומנות בעלת חשיבות מרכזית.

מסגרת מושגית

פיתוח מבחן "תנופה למחר" מבוסס על התמודדות ישירה עם סוגיות יסוד מרכזיות בתחום הערכת מיומנויות המאה ה-21. המבחן מציע גישות חדשניות שמספקות פתרונות ישימים לאתגרים תיאורטיים וטכנולוגיים, ומעמידות אותו בחזית המחקר העדכני.

הערכה הוליסטית של מיומנויות: המבחן מתמקד בהערכת מיומנויות קוגניטיביות כמכלול שזור ודינמי. הוא מתמודד עם אתגר ההפרדה המלאכותית ביניהן, וזאת מתוך הבנה כי בחיים האמיתיים נדרשת הפעלה סינרגטית של חשיבה ביקורתית, חשיבה יצירתית ואוריינות דיגיטלית. המשימות במבחן מדמות מצבים מורכבים המצריכים הפעלה משולבת של מיומנויות אלו.

הערכת ידע כתהליך אינטגרלי: המבחן מיישם גישה מתקדמת הרואה בידע תהליך אינטגרטיבי ודינמי. הוא בוחן כיצד התלמיד, באמצעות אוריינות מידע כעדשה מרכזית, משתמש במיומנויות באופן משולב כדי להפוך מידע גולמי לתוצר בעל ערך. תהליך זה מתחיל באיתור ובחירה של מקורות מידע רלוונטיים, ממשיך בארגון המידע ובשימוש בו, ומסתיים בייצוג והצגה של תוצר חדש. המבחן מעריך את השלבים השונים של התהליך ואת האופן שבו התלמיד מפעיל מלכה למעשה את החשיבה הביקורתית והיצירתית ואת האוריינות הדיגיטלית שלו.

הערכה של מיומנויות כלליות: המבחן מתמקד במיומנויות כלליות בתחומי ההומניסטיקה. המשימות בו אינן נשענות על ידע בתחום דעת ספציפי, אלא על הקשרים יום-יומיים ואוטנטיים שמאפשרים לבחון את המיומנות עצמה בנפרד מהידע. גישה זו מבטיחה שההערכה תתמקד ביכולות הקוגניטיביות הכלליות של התלמיד ותספק תמונה רחבה של יכולותיו.

הערכה המשלבת בינה מלאכותית: המבחן מכיר במציאות הטכנולוגית העכשווית ומשלב באופן מושכל את הבינה המלאכותית כחלק אינטגרלי מתהליך ההערכה. המבחן בוחן את יכולת התלמיד להשתמש בכלי זה באופן אוטונומי, אחראי וביקורתי כדי לפתור בעיות מורכבות, ובכך הוא הופך לכלי רלוונטי ועדכני לדור הנוכחי.

תוקף ומהימנות

ניתוח ראשוני של המחקר החלוץ, שכלל 934 תלמידים, הראה שהמבחן בעל מהימנות פנימית גבוהה (אלפא קרונברך גבוה מ-0.9) ומהימנות בין שופטים גבוהה (89.6% הסכמה). נעשו גם ניתוחי פריטים שבדקו את תפקודם, ויושמו מסקנות בנוגע לשיטת הציון של הפריטים. נבדקו הבדלים בביצועי המבחן בין קבוצות דמוגרפיות שונות ונמצא שאין הבדל בין קבוצות מגדר בביצועים, אך יש הבדל מובהק בין סוגי הפיקוח (עם יתרון לתלמידי הפיקוח הממלכתי) ובין רמות הרקע החברתי-כלכלי (עם יתרון לתלמידים בעלי רקע חברתי-כלכלי גבוה). כמו כן נמצא כי לא היה הבדל מובהק בביצועים בין מבחן באמצעות מחשב לבין מבחן באמצעות טלפון נייד.

נאספו עדויות ראשוניות לתוקף מתכנס ומבחין. נמצא כי ניסיון קודם בשימוש ביישומים דיגיטליים הוא מנבא מובהק להצלחה במבחן. כמו כן נבדק המתאם בין הציון הכולל במבחן לציונים שקיבלו אותם תלמידים במבחני תנופה בשפת אם ומדעים לכיתות ט'. נמצא כי המתאם של ציון המבחן עם תת-נושא לשון הוא $r=0.39$, בעוד המתאם של ציון המבחן עם תת-נושא חשיבה ביקורתית-יצירתית הוא $r=0.48$. ממצאים אלו הם עדות מסוימת לתוקף מתכנס (מדידה של חשיבה ביקורתית-יצירתית במבחני תנופה) ותוקף מבחין (מדידה של פריטי לשון במבחן תנופה בשפת אם).

השלבים הבאים

בעקבות מחקרי החלוץ והפילוסופים שקדמו לו נעשו שדרוגים במערכת, כגון פיתוח מודל התחברות חלופי, דף מוכנות ייעודי ושיפור ממשק המשתמש. זאת ועוד, התוצאות הראשוניות יורחבו באמצעות ניתוחי גורמים (Factor Analysis), שיסייעו להבין את הקשרים בין הפריטים השונים לבין צירי אוריינות המידע שנבדקו במבחן. ניתוחים אלו יסייעו גם בהתאמת הניסוחים ובשיור של הפריטים לתחומים הנבדקים במבחן. במקביל אנו מתקדמים לפיתוח מבחנים לשנים הבאות, שבהם נטמיע מסקנות על סמך ההתנסויות עד כה – לצד התאמות שיידרשו בשל התקדמות הטכנולוגיה.

לסיכום...

"תנופה למחר" הוא מבחן חדשני מסוגו, המותאם למסגרת התפיסתית של משרד החינוך ומעוגן במסגרת תיאורטית עדכנית. תהליך הפיתוח שלו היה שיטתי והתבסס על מתודולוגיה מחזורית של פיתוח, התנסות ושיפור מתמיד. בעקבותיהם נעשו שינויים בפרטי המבחן, במחווה ובאופן הבדיקה והציון. זאת נוסף על שדרוגים טכנולוגיים ושיפורים בממשק המשתמש, שנועדו להבטיח את יעילותו לקראת פריסה רחבה.

התוצאות הראשוניות של המבחן מבטיחות ביותר, בעיקר בהתחשב בעובדה שמדובר בפיתוח חדשני, והן מספקות בסיס איתן להעברה מדגמית בשנת תשפ"ו, לצד המשך דיוק ושיפור.

תהליך העבודה

תהליך הפיתוח של המבחן הוא מסע מדעי ורפלקטיבי, שכל שלב בו נועד לספק מידע קריטי שיזין את השלב הבא. הפיתוח התנהל במתכונת מחזורית של למידה, התנסות, ניתוח ושיפור מתמיד. להלן אבני הדרך המרכזיות – ובכלל זה אבני דרך בהחלטה, בפיתוח ובהטמעה.

פורסם מסמך דמות הבוגר , שהגדיר ארבע מיומנויות קוגניטיביות חוצות תחומים (חשיבה יצירתית, חשיבה ביקורתית, אוריינות מידע ואוריינות דיגיטלית) שאינן קשורות בהכרח לתחום דעת מוגדר.	3/2020
יום למידה משותף של המזכירות הפדגוגית ושל רמ"ה בנושא הערכת מיומנויות קוגניטיביות, בדגש על המיומנויות חוצות תחומים. במסגרתו מוצג גם מחקר פיזה בנושא חשיבה יצירתית.	12/2022
פורסמו מסקנות הצוות המשרדי שמינתה מנכ"לית המשרד לשם גיבוש מערך ההערכה החדש, ובהן נקבע שכחלק מהמערך יגובש מבחן חדש "במיומנויות במסגרת מקצועות המח"ר" (הדוח המלא).	3/2023
המבחן החדש מוכרז בחוזר מנכ"ל משרד החינוך מס' 0398, הקובע כי "רמ"ה תפתח מבדק דיגיטלי חדש הבודק מיומנויות וכישורים המתאימים למאה ה-21", שמיועד לתלמידי כיתות ו' וכיתות ט'.	7/2023
מפגש התייעצות ראשון , בהשתתפות צוות מגוון הכולל אנשי אקדמיה מובילים, פסיכומטריקאים ואנשי טכנולוגיה, נציגים מהמזכירות הפדגוגית, ותלמידים מאשדוד שבנו בכוחות עצמם AI. מבחן באמצעות AI.	9/2023
הושלם שלב הלמידה והמחקר , שארך כשלוש שנים, ובסופו פורסמה סקירת ספרות ומסגרת מושגית למדידה ("איך מודדים יצירתיות וביקורתיות? מסגרת מושגית ודוגמאות מהעולם").	3/2024
כינוס ראשון של ועדת ההיגוי לפיתוח המבחן, בראשות ד"ר ענת בן-סימון, לשעבר מנכ"לית מאל"ו, אנשי אקדמיה מובילים מהארץ ומהעולם, מורים ומדריכים. את הפיתוח מלווה ד"ר מירב אטאלי.	3/2024
ניסוי גישוש ראשוני בבית ספר בדרום הארץ, בהשתתפות 10 תלמידי כיתה ט'. נבדקה משימה שכללה שימוש בכלי בינה מלאכותית כדי להבין את תהליכי החשיבה של התלמידים.	6/2024

נקבעים תאריכי המבחן בחוזר מנכ"ל משרד החינוך מס' 0424: נקבע שהמבחן יועבר בפעם הראשונה בתשפ"ו, נוסף על המבחנים האחרים, ויתקיים ב-21/1/2026 לכיתות ט' וב-27/5/2026 לכיתות ו'.	9/2024
תחילת הפיתוח הטכנולוגי בשיתוף פעולה עם מינהל טכנולוגיות מידע במשרד, באמצעות חברת "סקילסט". הפיתוח תוכנן לאפשר היבחנות בסביבה פתוחה ושימוש במכשירי טלפון ניידים.	11/2024
השלמת הגרסה הראשונה של המבחן. ועדת ההיגוי מאשרת את הנוסח הראשון והוא מועבר במסגרת מיני מחקר חלוץ (פיילוט).	12/2024
השלמת התהליך של פיתוח המבחן. ועדת ההיגוי מאשרת את הגרסה המתוקנת של המבחן לקראת התנסות במחקר חלוץ רחב יחסית.	3/2025
מיני-מחקר חלוץ שני נערך בקרב 71 תלמידי כיתה ט': התהליך חשף קשיים טכניים אצל חלק מהתלמידים, והניע חשיבה על האורך והחזרתיות של המשימות. צוות הפיתוח הטמיע את ההמלצות.	3/2025
הצגה ראשונית במזה"פ ובמנה"פ. עם השלמת הנוסח הראשוני הוא הוצג במפגשים ייעודיים במזכירות הפדגוגית ובמינהל הפדגוגי, ונאספו התייחסויות ומחשבות לקראת ההמשך.	3/2025
התנעת תהליך ההטמעה. בפגישה של מנכ"ל משרד החינוך ומנכ"ל רמ"ה סוכם להתחיל בסבב הצגת המבחן לפורומים השונים במשרד, לקראת השקתו בשנת הלימודים תשפ"ו.	4/2025
פרסום מפרט המבחן. רמ"ה מעלה לאוויר את מפרט המבחן, לצד חומרים שיווקיים נוספים על אודות המבחן, המסגרת התפיסתית שלו וההיבטים התיאורטיים של הערכת המיומנויות.	5/2025
"תנופה למחר" מוצגת בפורומים שונים: הנהלת המשרד, מנהלי מחוזות, מפגשי פיקוח במחוזות, המזה"פ, המנה"פ, אגף אסטרטגיה, מינהל עובדי הוראה, מינהל טכנולוגיה וחדשנות ועוד.	5/2025
המחקר החלוץ. המבחן הועבר בפעם הראשונה בהיקף רחב בקרב כ-1,160 תלמידים מ-50 בתי ספר. מטרתו לבחון את איכות השאלות, תפקוד המערכת והתנהלות הבחינה.	6/2025
החלטה על היקף ההעברה: שר החינוך מקבל את המלצת רמ"ה ולפיה בשנה הראשונה יתקיים מבחן "תנופה למחר" במתכונת של "רוצה אני" (כך שבתי הספר יבחרו להשתתף בו), לצד מדגם ארצי מייצג.	7/2025

יסודות המבחן

ארבע המיומנויות הנבדקות

ההתפתחות המואצת של תחומי המידע והטכנולוגיה, ובעיקר הופעתה של הבינה המלאכותית היוצרת (Generative AI), משפיעה במידה ניכרת על דרכי הלמידה, ההוראה וההערכה. המידע נעשה זמין ונגיש בלחיצת כפתור, ותלמידים יכולים ליצור תוכן, לפתור בעיות וללמוד בכוחות עצמם באמצעות כלים טכנולוגיים מגוונים. נגישות זו מאפשרת למידה גמישה, יצירתית ומותאמת אישית, אך בד בבד מציבה אתגרים: הצפת מידע, קושי לזהות מקורות מהימנים, צורך לעבד מידע מורכב והתמודדות עם מידע משתנה במהירות.

לנוכח האתגרים, מערכת החינוך מתמקדת בפיתוח מספר מיומנויות קוגניטיביות חוצות תחומים:

אוריינות מידע	היכולת להגדיר את המידע הנדרש, לאתר, לארגן, להעריך ולהציג בדרך מושכלת.
חשיבה ביקורתית	היכולת לבחון ולהעריך מידע, דעות או רעיונות בדרך מושכלת, לגבש דעה או עמדה באופן עצמאי, לבחור בין חלופות ולקבל החלטות מנומקות.
חשיבה יצירתית	היכולת לחשוב על נושאים מוכרים בדרכים חדשות, להציע הסברים או פתרונות חלופיים, ליצור הקשרים חדשים משילוב של תחומי ידע שונים ולהפיק תוצרים מקוריים, רלוונטיים ובעלי ערך.
אוריינות דיגיטלית	היכולת להשתמש בטכנולוגיות מידע ותקשורת למגוון של צרכים ומטרות אישיות, קבוצתיות וחברתיות-לימודיות ובחיי היום-יום באופן הולם, יעיל ואחראי, ולהסתגל במהירות להתפתחויות ולשינויים טכנולוגיים.

במסמך "[דמות הבוגר](#)" מוצגים רכיבי הליבה והמיומנויות הקוגניטיביות בהלימה לספרות המקצועית, והם משקפים תהליכי הוראה, למידה ופיתוח חשיבה לאורך זמן. עם זאת, בהערכת מיומנויות קוגניטיביות יש מורכבות הנובעת מהאופי תלוי-ההקשר של המיומנויות, מהחפיפה הטבעית ביניהן, מהעובדה שהן נוגעות לתחומים נרחבים, וכן מהצורך להעריך בסביבה פתוחה שיש בה מקורות מידע וכלים טכנולוגיים זמינים ללומד. משום כך המבחן בודק את המיומנויות כשילוב אינטגרטיבי, בתוך משימות המדמות את המציאות.

גישה זו הכרחית גם מכיוון שכלל המיומנויות הקוגניטיביות המתוארות לעיל נשענות על בסיס מוצק של [אוריינות לשונית](#). היכולת להבין, לפרש ולנסח רעיונות בשפה דבורה או כתובה היא תנאי הכרחי לשם פיתוח ויישום של חשיבה ביקורתית, אוריינות מידע, חשיבה יצירתית ואוריינות דיגיטלית. לכן המבחן, המעריך את כל המיומנויות הללו, רואה באוריינות הלשונית מיומנות יסוד המשתלבת באופן בלתי נפרד בכל שלבי ההערכה.

אוריינות מידע כעדשה להערכה

במבחן "תנופה למחר" נבחרה אוריינות המידע לשמש עדשה מרכזית להערכה, שדרכה אפשר להתבונן גם בהיבטים של חשיבה ביקורתית, חשיבה יצירתית ואוריינות דיגיטלית – לאורך תהליך שלם של עבודה עם מידע, משלב האיתור ועד ההצגה. הבחירה בעדשה זו נועדה לשקף כיצד מיומנויות אלו באות לידי ביטוי בפועל, והיא נשענת על שיקולים פדגוגיים, תיאורטיים ומעשיים. עם זאת, כל אחת מהמיומנויות יכולה לשמש גם מוקד עצמאי של הוראה או הערכה, כלומר עדשה ממוקדת לבחינה מעמיקה של מיומנות אחת. הבחירה בעדשה תלויה בהקשר ובמטרות ההערכה: האם המיקוד הוא בהעמקה בתחום ספציפי, או באינטגרציה של מיומנויות בתוך משימה מורכבת ורב-שלבית.

המפגש בין חשיבה ביקורתית, חשיבה יצירתית ואוריינות דיגיטליות



חשיבה יצירתית

היכולת לחשוב באופן חדשני וליצור רעיונות או פתרונות חדשים.



חשיבה ביקורתית

היכולת לבחון ולהעריך מידע, דעות או רעיונות באופן מושכל, לגבש עמדה ולבחור בין חלופות.



אוריינות דיגיטלית

היכולת להשתמש בטכנולוגיות מידע ותקשורת למגוון צרכים ומטרות. באופן הולם ויעיל תוך שמירה על אתיקה ואחריות.



אוריינות מידע

אוריינות לשונית

תפיסות יסוד

המבחן החדש מבוסס על שלושה יסודות עיקריים: מיומנויות, חדשנות ורלוונטיות. כל אחד מהם משקף את הגישה מעשית של המבחן, שנועד להעריך את יכולות התלמידים בסיטואציות אמיתיות מהחיים תוך שימוש בטכנולוגיות ובכלים חדשניים.

מיומנויות – לא רק לדעת, אלא גם לדעת לעשות: המבחן מעריך את אוריינות המידע של הלומדים, ובדק כיצד באות בה לידי ביטוי חשיבה ביקורתית, חשיבה יצירתית ואוריינות דיגיטלית – לאורך שלבי העבודה עם מידע, ובהתאם לאתגרי הסביבה הדיגיטלית הפתוחה.

חדשנות – טכנולוגיה בשירות ההערכה: החדשנות במבחן באה לידי ביטוי באופיו הפתוח ובשילוב טכנולוגיות מתקדמות כחלק בלתי נפרד מתהליך ההיבחנות. התלמידים נבחנים באמצעות אמצעי קצה אישיים, כגון טלפון נייד או מחשב, ובמהלך ההיבחנות הם רשאים להיעזר גם בכלים דיגיטליים מגוונים, כגון מנועי חיפוש וכלי בינה מלאכותית. הסביבה הדיגיטלית מאפשרת לאסוף, לעבד ולבחון מידע בזמן אמת, ומעודדת חקר עצמאי, חשיבה ביקורתית ויצירתיות באופן המדמה תהליכי למידה מודרניים.

רלוונטיות – כשההערכה פוגשת את החיים: הרלוונטיות של המבחן נובעת מהקשר הישיר בין משימות ההערכה לבין עולמם של הלומדים. המשימות מדמות סיטואציות מחיי היום-יום של התלמידים, ומתמקדות באתגרים ממשיים: התמודדות עם מידע עודף, קבלת החלטות בתנאים משתנים ושימוש מושכל בטכנולוגיה.

עקרונות הפיתוח

לנוכח המורכבות שבהערכת מיומנויות קוגניטיביות בסביבה דיגיטלית פתוחה, גובשו מספר עקרונות ששימשו בסיס לפיתוח המבחן:

הערכה בתוך הקשר: מטרתה לבדוק כיצד התלמיד מתפקד בתוך הקשר המדמה את העולם האמיתי. המיקוד הוא ביכולתו להשתמש בידע, להפעיל שיקול דעת ולבחור אסטרטגיות לפעולה.

סביבה פתוחה היא המציאות: ההערכה נעשית בסביבה פתוחה, המדמה את המציאות שבה פועלים התלמידים באמצעות אמצעי קצה אישיים (כגון טלפון נייד או מחשב), תוך אפשרות להיעזר בכלים דיגיטליים מגוונים, ובהם כלי בינה מלאכותית. שימוש נבון בכלים אלו הוא חלק בלתי נפרד מאוריינות המידע בעידן הנוכחי.

תפקוד משולב: ההתמודדות עם משימות במציאות מצריכה שילוב של מיומנויות מסוגים שונים. לכן המדידה שואפת לזהות ולתאר את השילוב הזה ולא לפרק את הביצוע למיומנויות נפרדות.

מדידה רב-ממדית: ההערכה שואפת לזהות ולתאר פרופיל יכולות מגוון, המבליט תחומי חוזק ואתגר ומאפשר להתרשם כיצד התלמידים פועלים.

המסגרת המושגית

סוגיות יסוד בהערכת מיומנויות במאה ה-21

מבחן "תנופה למחר" הוא כלי הערכה חדשני, המבוסס על אינטגרציה יישומית של עקרונות תיאורטיים מובילים מתחום הערכת מיומנויות המאה ה-21. פיתוחו הצריך חיבור הדוק לתיאוריות העדכניות ביותר והבנה מעמיקה של הצרכים העכשוויים. פרק זה בא לשקף את התיאוריות והמחקרים שעיצבו את המבחן, ואת סוגיות היסוד שאיתן התמודדנו בתהליך.

ככלל, בתהליך הפיתוח התמודדנו עם ארבע סוגיות יסוד אשר בהן נדרשנו לשנות את המודלים המסורתיים ששימשו את רמ"ה במדידת הידע לאורך כעשרים שנות קיומה – ולהתאימם למדידה של כישורי המאה ה-21: מעבר מהערכת מיומנויות נפרדות להערכה הוליסטית הרואה בהן מכלול שזור וסינרגטי; התייחסות לממד הידע כאל רצף אינטגרטיבי, שמתחיל בידע כמקור (איסוף וסינון מידע) ומסתיים בידע כתוצר (יצירת ידע חדש), ובכך הופך את הלומד מצרכן מידע ליוצר ידע פעיל; מעבר מהערכת מיומנויות המבוססות על ידע ספציפי בתחום להערכה המתמקדת במיומנויות כלליות וחוצות תחומים שאינן תלויות בידע קודם; ושילוב מושכל של בינה מלאכותית (AI) בתהליך ההערכה, כחלק מהמגמה העולמית של מעבר ממבחנים המתעלמים מכלי זה למבחנים שמשלבים אותו כחלק אינטגרלי מהפתרון.

בכל ארבע הסוגיות נדרשנו לקבל החלטות לצורך פיתוח המבחן, אשר עלו בקנה אחד עם גישות מודרניות שונות. כך, המבחן התגבר על המחלוקות המחקריות המסורתיות בנוגע לבידוד מיומנויות, לתלות הבלעדית בידע קודם ולגישה של הערכה בסביבה טכנולוגית סגורה, וזאת כדי לספק תמונה מלאה יותר של יכולותיו הקוגניטיביות של התלמיד.

1. מהערכת חלקים להערכת השלם

בעידן שבו הידע זמין ונגיש בקלות, מערכות החינוך חייבות להעביר את הדגש משינון מידע לפיתוח מיומנויות חשיבה (OECD, 2017). עם המעבר לדגש זה על פיתוח מיומנויות עולה השאלה כיצד נכון להעריך אותן. אחד האתגרים המרכזיים בהערכת מיומנויות המאה ה-21 הוא אופן המדידה שלהן: האם יש לראות בכל אחת מהמיומנויות יחידה נפרדת הנמדדת בפני עצמה, או שמא נכון לראות בהן מכלול שזור הפועל בסינרגיה כחלק מתהליך קוגניטיבי מורכב?

הערכה של מיומנויות כיחידות נפרדות: גישה זו, שהייתה נפוצה יותר בעבר, מתמקדת בהערכת כל מיומנות בנפרד. כך, למשל, נבנו מבחנים ייעודיים לחשיבה ביקורתית, מבחנים נפרדים ליצירתיות, וכלי הערכה אחרים לאוריינות דיגיטלית. היתרון בגישה זו הוא האפשרות לקבוע ציון מדויק לכל מיומנות בנפרד, מתוך הנחה שאפשר לבודד אותה לצורך מדידה.

הערכה של מיומנויות כמכלול: שלא כגישה הקודמת, המתמקדת במדידת כל מיומנות בנפרד, הגישה הנפוצה כיום במחקר ובמסגרות בין-לאומיות רואה במיומנויות מכלול שזור ודינמי. היא נשענת על ההבנה כי בחיים האמיתיים משימות מורכבות מצריכות הפעלה משולבת וסינרגית של מיומנויות מגוונות (Weldon, 2019). בהתאם לזה, מחקרים שונים דוגמת PISA מעריכים מיומנויות בהקשרים רחבים המדמים מציאות מורכבת, ולא כסך חלקים מבודדים. לדוגמה: פתרון

בעיות שיתופי (OECD, 2015). גישה זו נתמכת גם בעבודותיו של סטרנברג (Sternberg 2024). המדגיש כי הצלחה בעולם האמיתי נשענת על שילוב בין יכולות אנליטיות, יצירתיות ופרקטיות – כחלק מתפיסת "האינטליגנציה האדפטיבית", המוגדרת כיכולת להתאים את עצמך לסביבה משתנה כדי להשיג מטרות.

השינוי התפיסתי מגישת החלקים לגישת השלם בא לידי ביטוי בכמה מיומנויות. לדוגמה:

א. אוריינות מידע ואוריינות דיגיטלית: ההתפתחות של האוריינות הדיגיטלית מדגימה באופן מוחשי את המעבר מגישת החלקים לגישת השלם. פרופסור מרקוס לינינג (Leaning, 2019), מומחה לחינוך מדיה דיגיטלית, סוקר כיצד שני מושגים מרכזיים התפתחו במקביל, אך בנפרד זה מזה, וכל אחד מהם התייחס רק לחלק מהתמונה. אוריינות מידע התמקדה בעיקר בהיבטים טכניים-פונקציונליים, כגון היכולת לאתר, לארגן ולסכם מידע באופן יעיל. אולם היא חסרה התייחסות ביקורתית למידע. מנגד, אוריינות מדיה סיפקה את הממד הביקורתי החיוני, והתמקדה בניתוח של תכנים ומסרים תרבותיים וחברתיים. עם זאת היא לא התייחסה לממד הטכנולוגי החדש, ולא נתנה מענה הולם לאופן שבו הטכנולוגיה הדיגיטלית עצמה – כגון רשתות חברתיות ובוטס – משפיעה על המידע ומעצבת אותו. לכן לינינג מציע לראות באוריינות דיגיטלית "מטרייה" (Umbrella Term) המאחדת את שני התחומים ומציבה את "האוריינות הדיגיטלית הביקורתית" במרכז. בגישה זו, האדם אינו רק אוסף ומארגן מידע, אלא גם בוחן באופן ביקורתי את המדיום והפלטפורמה שבאמצעותם המידע מגיע, תוך הבנה כיצד הטכנולוגיה עצמה משפיעה על התוכן. גישה זו מכלילה את הכישורים הטכניים של אוריינות המידע עם החשיבה הביקורתית של אוריינות המדיה, ובכך היא מאפשרת למשתמש להבין לא רק את התוכן אלא גם את המדיום הטכנולוגי המתווך אותו, ולפעול עליהם. בדרך זו המשתמש יכול להפוך מ"צרכן מידע" פסיבי ל"יוצר ידע" ביקורתי ופעיל.

ב. חשיבה ביקורתית וחשיבה יצירתית: גם הקשר בין חשיבה ביקורתית לחשיבה יצירתית ממחיש את הצורך בגישה אינטגרטיבית. [המסגרת המושגית](#) שהונחה כבסיס למבחן במיומנויות חשיבה ביקורתית ויצירתית מדגישה כי יש קשר הדוק ותלות הדדית בין שתי המיומנויות. במקום לראות בהן ישויות נפרדות, ההבנה היא שאין תהליך – יצירתי או ביקורתי – שאינו מחייב רכיבים משתי המיומנויות גם יחד. לדוגמה, חשיבה יצירתית מצריכה רכיבי הערכה וניתוח (חשיבה ביקורתית) כדי לוודא שרעיונות חדשים ישימים ובעלי ערך. מנגד, חשיבה ביקורתית מצריכה חשיבה יצירתית כדי להעלות חלופות שונות לפתרון בעיות. המסמך מציג את התלות ההדדית בין שתי המיומנויות גם בהקשר של חשיבה מתכנסת ומסתעפת: חשיבה ביקורתית היא בעיקרה מתכנסת, ואילו חשיבה יצירתית היא מסתעפת. אולם ההבנה היא ששתיהן כאחד הן אבני יסוד של החשיבה האנושית, ושתיהן נדרשות במידה זו או אחרת בכל פעולה קוגניטיבית מורכבת, מה שמוביל להמלצה לפתח שאלות מבחן המשלבות את שתי המיומנויות במקביל.

הביטוי לגישה זו במבחן "תנופה למחר": בהתאם למגמות מחקריות עדכניות, מבחן "תנופה למחר" מתבסס על הגישה הרואה במיומנויות מכלול שזור הפועל בסינרגיה ולא אוסף של יכולות נפרדות זו מזו. המשימות במבחן מעוצבות כהדמיות של מצבים אותנטיים מהחיים האמיתיים, שבהם הלומד נדרש לשלב בין מיומנויות שונות בתוך הקשר אחד שלם. ההערכה מתמקדת לא ביכולת להפעיל כל מיומנות בפני עצמה, אלא ביכולת להפעילן יחד באופן אינטגרטיבי – כפי שנדרש במצבים מורכבים בעולם המשתנה.

2. אוריינות מידע כמסגרת לארגון המיומנויות

במענה לאתגר הערכת השלם, **אוריינות המידע** מדגימה גישה אינטגרטיבית, שכן היא רואה במיומנויות לא אלמנטים נפרדים אלא חלק מתהליך דינמי. מסגרת זו מתמקדת ביחסו של הלומד למידע – הן כמשאב שיש לברור, לארגן ולהבין (**ידע כמקור**), והן כתוצר חדש הנולד מתוך עיבוד, סינתזה וארגון מחדש (**ידע כתוצר**). התפיסה היא שהמיומנויות אינן פועלות בנפרד, אלא הן שזורות זו בזו באופן אינטגרלי. לפיכך, הפרדה ביניהן לצורכי הערכה היא עניין מורכב ואתגר מרכזי בהערכת מיומנויות הנדרשות באוריינות מידע (Ananiadou & Claro, 2009).

ידע כמקור (Information as a Source) מתייחס ליכולת של הלומד לאתר, לסנן, להעריך ולפרש מידע ממקורות שונים. מיומנות זו מחייבת חשיבה ביקורתית, הבחנה בין עובדות לדעות, הערכת מהימנותם של מקורות וזיהוי של הטיות או מניפולציות. בהקשר זה, המידע נתפס כמשאב בסיסי שיש לגשת אליו באופן מושכל ואסטרטגי – כך שהתלמיד יידע לבחור מקורות רלוונטיים ואמינים, ולאמץ פרקטיקות של אוריינות מידע ודיגיטל החיוניות לתפקוד מיטבי בעידן הנוכחי. לעומת זאת, ידע כתוצר (Information as a Product) מתמקד ביכולתו של הלומד להפוך מידע גולמי לידע חדש, רלוונטי ובעל ערך. מדובר בתהליך של עיבוד, ארגון, סינתזה והצגה של מידע, המאפשר לגבש תובנות, להסיק מסקנות ולפתח תוצרים כגון דוחות, מצגות, פתרונות חדשניים או רעיונות מקוריים. כאן הדגש עובר מהיות הלומד "צרכן מידע" בלבד ליכולתו לפעול כיוצר ידע פעיל (Active Knowledge Producer).

המסגרת המושגית הזאת משתלבת היטב עם הגדרת אוריינות המידע בדמות הבוגר, הכוללת את הרכיבים האלה: **איתור מידע** (להבין איזה מידע נדרש, להכיר מנועי חיפוש, להכיר כיצד הידע מאורגן, לאתר מידע ביעילות); **הערכת מידע** (להשוות בין מקורות מידע, להעריך מידע באופן ביקורתי); **ארגון מידע** (למזג בין מקורות מידע, לארגן, למיין ולתעד מידע); **שימוש במידע** (לסכם ולעבד מידע, לבחון השערות ולהסיק מסקנות, לעשות שימוש מוסרי במידע); **וייצוג מידע** (לייצג ולהציג מידע בדרכים שונות ובהתאם למטרה). רכיבי המיומנות "אוריינות מידע", או תת-המיומנויות הללו, מסודרים כשלבים עוקבים בתהליך מלא שתחילתו באיסוף מידע וסופו בתוצר.

כפי שאפשר לראות להלן, בכל אחד מהשלבים של עיבוד המידע משלב באופן סינרגטי את אוריינות המידע לצד חשיבה ביקורתית ויצירתית, וכל המיומנויות הללו פועלות יחד באופן שווה ומשלימות זו את זו.

איתור מידע והערכת מידע: פעולת האיתור מלווה מיד בפעולת הערכה, שכן תהליך הסינון והבקרה מתחילים כבר מרגע החיפוש. כאשר לומד מחפש מידע הוא גם מפעיל שיקול דעת ביקורתי כדי לבחור את המקורות המתאימים ביותר. תהליך זה מערב את החשיבה הביקורתית ככלי מרכזי לסינון ואימות המידע, המאפשר להבחין בין טענות לעובדות, לזהות הנחות סמויות, להעריך את אמינות המקור ולברור את המידע הנחוץ מתוך ים הנתונים הנגיש. בד בבד נדרשת חשיבה יצירתית כדי לפתח אסטרטגיות חיפוש חדשניות ולמצוא קשרים חדשים בין נתונים ממקורות שונים. שתי מיומנויות חשיבה אלו פועלות בתוך מסגרת של אוריינות דיגיטלית, המאפשרת למשתמש לבחור את כלי החיפוש הנכון ולהשתמש בו ביעילות, תוך הבנת ההבדלים וההטיות הקיימות בין הפלטפורמות. נוסף על כך, האוריינות הדיגיטלית מאפשרת לנתח את מאפייני המדיום, למשל – לזהות תוכן שנוצר על ידי בינה מלאכותית, ובכך היא מחזקת את החשיבה הביקורתית.

ארגון מידע ושימוש בו: השלב של ארגון המידע והשימוש בו מצריך חשיבה ביקורתית כדי למיין אותו באופן לוגי, לנתח את הנתונים ולגבש תובנות. עם זאת, בשלב זה נדרשת גם חשיבה יצירתית כדי למזג מידע ממקורות שונים וליצור תמונה שלמה וקוהרנטית. תהליך זה נעשה באמצעות אוריינות דיגיטלית, המאפשרת שימוש בכלים טכנולוגיים מתקדמים כגון מפות חשיבה דיגיטליות וטבלאות, שהם הכרחיים לארגון ולעיבוד של מידע בהיקף רחב. כל אלה הופכים את הלומד מצרכן מידע ליוצר ידע פעיל, המסוגל להשתמש במידע כדי להסיק מסקנות ולפתור בעיות באופן מעמיק.

ייצוג מידע והצגתו: הייצוג וההצגה הם שלב נוסף וקריטי בתהליך עיבוד הידע, המביא לידי ביטוי את שילובן הסינרגטי של המיומנויות. החשיבה היצירתית מאפשרת ללומד לחשוב "מחוץ לקופסה" ולבחור דרך חדשנית ומקורית להצגת המידע – אם בסיפור, באינפוגרפיקה או בסרטון. בה בעת החשיבה הביקורתית חיונית כדי לוודא שהייצוג הסופי לא רק יצירתי, אלא גם מתאים למטרה, לקהל היעד ולאמינות המקורות. היא מאפשרת לבחון את הדיוק, הבהירות והאפקטיביות של המסר. כל אלה מתאפשרים באמצעות אוריינות דיגיטלית, המעניקה את הכלים הטכנולוגיים להפקת תוצרים מגוונים ומתקדמים כגון מצגות אינטראקטיביות, דוחות מקצועיים או סרטוני אנימציה. בכך הלומד הופך ממעבד מידע למי שיוצר ידע ומוסר אותו באופן אפקטיבי.

הביטוי לגישה זו במבחן "תנופה למחר": מבחן "תנופה למחר" מיישם הלכה למעשה את התפיסה האינטגרטיבית כלפי ידע, באמצעות שימוש באוריינות מידע כ"עדשה" מאחדת. במקום לבחון כל אחת מהמיומנויות בנפרד, המבחן מעריך כיצד הן פועלות יחד באופן סינרגטי בתהליכי עבודה עם מידע. בהיבט זה, משימות המבחן אותנטיות ומדמות מצבים מורכבים מהחיים, כגון הצורך לחפש מידע במרחב הדיגיטלי, לסנן מקורות ולהעריך את מהימנותם. בדרך זו המבחן בודק את התהליך הכולל שבו תלמיד הופך ידע גולמי לידע חדש, רלוונטי ובעל ערך. יכולת זו, המדגימה מעבר מ"צרכן מידע" ל"יוצר ידע פעיל", היא לב ליבו של המבחן, והיא באה לידי ביטוי ביכולתו של הלומד להשתמש במיומנויות באופן אינטגרטיבי כדי להסיק מסקנות ולפתור בעיות באופן מעמיק.

3. הערכה המבוססת על ידע כללי – כמו בחיים

סוגיה מרכזית בהערכת מיומנויות היא השאלה אם יש להישען על ידע תחומי קודם, או שמא יש להתרכז במיומנויות חוצות תחומים שאינן תלויות בתחום דעת ספציפי. ההתפתחות במבחני PISA (Greiff, Holt & Funke, 2013) ממחישה שתי גישות עיקריות הבוחנות שאלה זו:

פתרון בעיות אנליטי (Analytical Problem Solving): גישה מסורתית המכונה גם "פתרון בעיות ספציפיות לתחום" (Domain-specific problem-solving skills). היא מבוססת על ההנחה כי ידע בתחום הדעת הוא תנאי מקדים והכרחי להפעלת מיומנויות. לפי גישה זו, ידע תחומי משמש "ידע חוקים" (rule knowledge) ובסיס שעליו נבנות המיומנויות, כך שכדי לבצע משימות מורכבות בתחום מסוים נדרשת שליטה בידע הרלוונטי. לדוגמה, כדי להפעיל חשיבה ביקורתית על מאמר בפיזיקה יש לשלוט בעקרונות הפיזיקליים, וכדי לנתח תהליך כימי יש להכיר היטב את עקרונות הכימיה. במסגרת גישה זו, ההערכה התרכזה בעיקר בתוכן הלימודי הספציפי של כל תחום דעת, ופחות במיומנויות חשיבה כלליות ושימושים שאינן קשורות לתוכן ספציפי (Schäfer, Reuter, Karbach, & Leuchter, 2024). לדוגמה: מבחן PISA 2003 העריך את יכולתם של התלמידים לפתור בעיות הקשורות קשר הדוק לתחומי דעת ספציפיים, כגון מתמטיקה ומדעים. המשימות היו סטטיות ולא אינטראקטיביות, ובוצעו בשיטת נייר ועיפרון. המבחן לא העריך

את יכולת התלמיד לגלות מידע חדש או ללמוד על מערכת בלתי מוכרת מתוך ניסוי וטעייה. במקום זאת הוא התמקד בעיקר במיומנויות בתוך מסגרת ידע סגורה (Greiff, Holt & Funke, 2013).

פתרון בעיות אינטראקטיבי (Interactive Problem Solving): גישת פתרון הבעיות האינטראקטיבי (Interactive Problem Solving) מייצגת התפתחות חשובה בתפיסה ובהערכה של מיומנויות חשיבה, והיא הגישה המובילה את תפיסת ה-OECD בשנים האחרונות. בניגוד לגישה המבוססת על ידע תוכן, גישה זו מתמקדת במיומנויות כלליות ובלתי תלויות בתחום דעת ספציפי (Domain-independent skills). היא נועדה לשקף אתגרים יום-יומיים ואת המורכבות של העולם המודרני, שבו המידע הנדרש לפתרון בעיה אינו תמיד זמין במלואו מההתחלה (Greiff, Holt & Funke, 2013). גישה זו רלוונטית במיוחד למציאות המשתנה של המאה ה-21, ושואפת למדוד את המיומנות עצמה ולא את רמת הידע התוכני שברקע. היא מומחשת בבירור במבחני PISA 2012, שנערכו בסביבה ממוחשבת ודרשו מהנבחן לאסוף מידע באופן פעיל באמצעות אינטראקציה דינמית, בניגוד לבעיות אנליטיות שבהן כל המידע זמין מראש. כותבי המאמר מציינים כי נוסף על פתרון בעיות אנליטי, פתרון בעיות עשוי לכלול תהליכים מנטליים כלליים מורכבים שאינם מוגבלים לתחומי דעת ספציפיים.

לסיכום, בעוד פתרון בעיות אנליטי ואינטראקטיבי נראים שונים זה מזה, הם למעשה פנים שונות של אותה מסגרת של פתרון בעיות. ההתפתחות ממבחן פיזה 2003 למבחן פיזה 2012 מדגימה את המעבר מהתמקדות בשאלות המעריכות מיומנויות על סמך ידע קודם בתחום דעת, לשאלות המעריכות מיומנויות על בסיס ידע גנרי שאינו בהכרח מקושר לתחום דעת ספציפי. **שתי הגישות אינן סותרות זו את זו, אלא הן היבטים שונים של ההערכה.** יתרה מזאת, שתי הגישות נחוצות ומשלימות זו את זו. למשל, כאשר תלמיד נבחן בשאלה שמצריכה חשיבה ביקורתית המבוססת על ידע במדעים, הקושי יכול לנבוע גם מהיעדר ידע מספיק בתחום הדעת, ואין זה מעיד בהכרח על יכולת החשיבה הביקורתית שלו בהקשרים אחרים. גישת פתרון הבעיות האינטראקטיבי מציגה היבט נוסף וחשוב, שכן היא מאפשרת למדוד את המיומנות הכללית בנפרד מהתלות בידע ספציפי, ובכך מספקת תמונה מלאה יותר של יכולותיו הקוגניטיביות של התלמיד.

סוגיית התוקף בהערכת מיומנויות היא, כמובן, מורכבת. האתגר הוא להבטיח שהמבחן אכן מודד את יכולת המטרה בלבד וכי ביצועי התלמיד אינם מושפעים מיכולות אחרות. מבחן המעריך מיומנויות על בסיס של ידע בתחום דעת (הגישה האנליטית) נתקל בבעיה מבנית של תוקף, שהרי קושי של תלמיד יכול לנבוע הן מחולשה במיומנות והן מחוסר ידע בתחום הדעת. אלא שיש נקודת תורפה גם בגישה האינטראקטיבית (המעריכה מיומנויות כלליות כגון חשיבה ביקורתית, ללא תלות בידע קודם): כאן הקושי של התלמיד יכול לנבוע הן מחולשה במיומנות החשיבה הביקורתית והן מחולשה ביכולות קוגניטיביות אחרות כגון זיכרון, שפה ומהירות עיבוד. במילים אחרות: בעוד מבחן שמעריך ידע בתחום דעת מסוים אינו מתמודד עם בעיות מבניות, מבחן שמעריך מיומנויות מתמודד עם אתגר של תוקף בכל דרך שבה ייעשה – אם במסגרת של תחום תוכן מוגדר (כגון מדעים) ואם מחוצה לה.

הביטוי לסוגיה זו במבחן "תנופה למחר": מבחן "תנופה למחר" מיישם הלכה למעשה את העקרונות של הגישה שחוצה תחומים. המשימות במבחן אינן נשענות על ידע דיסציפלינרי צר, אלא על מצבים אותנטיים מחיי היום-יום של הלומדים, המצריכים ידע כללי. כך, למשל, התלמידים נדרשים להעריך מהימנות של ידיעות ברשתות החברתיות, לתכנן פעילות, או לבחור מקורות

מידע לצורך קבלת החלטות. משימות אלו מדמות מצבים בלתי־שגרתיים שבהם המידע חלקי, משתנה ולעיתים אף מטעה – בדומה לאתגרי העולם הדינמי שבו אנו חיים.

4. מהערכה שמתעלמת מ-AI להערכה שמשלבת AI

בעידן שבו כלים דיגיטליים ובינה מלאכותית (AI) נגישים וקלים להשגה, צפה סוגיה מרכזית בדבר האופן שבו אנו מעריכים מיומנויות. התפיסה של OECD 2030 מכירה בעובדה שכיום גם המשימות המורכבות ביותר אינן מצריכות בהכרח שינון מידע, אלא יכולת להשתמש במגוון רחב של כלים ומשאבים כדי לפתור בעיות מורכבות. בהתאם לתפיסה זו, במקום להעריך "האם התלמיד יודע את התשובה?" מעריכים "האם התלמיד יודע כיצד להשתמש בכלים שעומדים לרשותו כדי להגיע לפתרון?"

לודג' ועמיתיו (Lodge, Howard & Broadbent, 2023) מציג טקסונומיה של שלוש רמות המדרגת את רמת המבחן בהתאם לאינטגרציה שלו עם הבינה המלאכותית:

- **רמה 1 – המבחן ללא בינה מלאכותית (The AI-less Exam):** גישת ההערכה בסביבה סגורה, שבה המבחן מתעלם מקיומו של ה-AI והוא נתפס ככלי פסול. מטרת ההערכה היא לבחון יכולות "טהורות" שרכש התלמיד, במנותק מכל סיוע חיצוני.
- **רמה 2 – המבחן שמשלב בינה מלאכותית (The AI-inclusive Exam):** גישת ההערכה בסביבה פתוחה, שבה ה-AI הוא כלי מוכר ולגיטימי לשימוש. ההערכה מתמקדת בשימוש האוטונומי והמושכל של התלמיד בכלים דיגיטליים, והמבחן מודד את יכולתו להשתמש בהם כדי לפתור בעיות מורכבות.
- **רמה 3 – המבחן עם בינה מלאכותית (The AI-as-an-equal Exam):** גישה מתקדמת יותר, שבה ה-AI הוא חלק בלתי נפרד מליבת המבחן. במודל זה התלמיד מתבקש לעבוד בשיתוף פעולה עם המכונה כדי לבצע משימות ספציפיות, ובדרך זו נבחנת יכולתו לפתח ולנתח פלט מתוכנת.

הביטוי לנושא זה במבחן "תנופה למחר": המבחן מיישם הלכה למעשה את המעבר לשילוב מושכל של AI, ובגמישות בין רמות 2 ו-3 בטקסונומיה של לודג'. מצד אחד הוא רואה ב-AI כלי לגיטימי לפתרון בעיות מורכבות (רמה 2), ובוחן את יכולתו של התלמיד להשתמש בו באופן מושכל ואוטונומי. מצד אחר, הוא מאפשר לתלמידים להרחיב את האינטראקציה עם הכלי – לפתח רצף הנחיות מורכבות, להטיל ספק במגבלות ה-AI, ולשלב אותו כשותף ביקורתי בתהליך הפתרון (רמה 3).

באופן זה המבחן אינו בוחן רק את הידע או את התוצר הסופי, אלא גם את הבחירות האסטרטגיות של התלמיד בכל הנוגע ל-AI – באיזו מידה הוא יודע להפעיל שיקול דעת, אחריות וביקורתיות בשימוש בו. גישה זו הופכת את המבחן לרלוונטי במיוחד בעולם שבו היכולת להשתמש בכלי אינה טכנית בלבד, אלא היא חלק מהותי מהמיומנויות הדרושות להתמודדות עם אתגרי המאה ה-21.

מבט לעתיד: מבחן מתפתח בעולם משתנה

מבחן "תנופה למחר" הוא כלי הערכה יישומי המשלב גישות תיאוריות מובילות כדי למדוד מיומנויות עתידיות באופן הוליסטי. הוא מתגבר על המחלוקות המחקריות בנוגע לבידוד מיומנויות, לתלות הבלעדית בידע קודם ולגישה המיושנת של הערכה בסביבה סגורה. באמצעות הצבת המיומנויות בתוך הקשרים אותנטיים ויום-יומיים, המבחן לא רק מודד את היכולות הנדרשות במאה ה-21, אלא גם מעביר מסר חינוכי בדבר חשיבותן ובדבר הצורך לטפחן בבית הספר באופן רחבי.

ואולם, אף אם המבחן נשען על יסודות מוצקים ומוכרים, בשל עצם היותו חלוצי הוא עדיין מפלס דרכו בעולם משתנה של אי-ודאות. מכאן שהמבחן צפוי להשתנות ולהתעצב משנה לשנה, לצד התפתחות הטכנולוגיה והבינה המלאכותית. אנו צופים אתגרים חדשים כל הזמן, ומכאן שנראה כי לעולם לא נוכל להיות "מוכנים לחלוטין". המבחן נולד מתוך ההבנה שהטכנולוגיה מתפתחת במהירות, ולכן כל מודל הערכה של מיומנויות חוצות תחומים עלול להיות מיושן עוד בטרם יושק. בהתאם לזה אנו רואים במבחן זה נקודת התחלה חשובה, ולא יעד סופי. הממצאים שיופקו ממנו והתקדמות הטכנולוגיה בעתיד יסייעו לנו לחדד, לעדכן ולשפר אותו בהתמדה, כדי שימשיך להיות רלוונטי ויעיל גם לדורות הבאים.

תהליך הפיתוח

שלבי הפיתוח

פיתוח המבחן החל לפני כשלוש שנים, על בסיס תהליך מקצועי ומתמשך של למידה ומחקר. מדובר במהלך שיטתי ומבוסס על ראיות, שכלל שלבים רבים כפי שיפורט בהמשך. המהלך לווה ביעוץ של מומחים מהאקדמיה ונציגי משרד החינוך, והוא מתנהל במסגרת ועדות היגוי ייעודיות. התהליך כלל פיתוח משימות, התנסות בשטח, ניתוחים סטטיסטיים מדוקדקים ודיוק מתמשך של הפריטים והמחווה, כל אלה נועדו להבטיח את איכותו המקצועית ולוודא כי הוא משקף באופן הוגן את יכולות התלמידים. להלן ששת השלבים העיקריים בתהליך הפיתוח:

1. תהליך למידה בתוך רמ"ה

תהליך הפיתוח החל באגף פיתוח כלי הערכה, והתמקד בלמידה מעמיקה של האתגר מנקודות מבט שונות. בשלב זה פוצל האתגר לתת-נושאים, וצוות הפיתוח בחן לעומק את הנעשה בעולם. הלמידה התמקדה בכמה תחומים מרכזיים:

הערכת מיומנויות: כחלק מתהליך הלמידה נעשתה [סקירת ספרות](#) מקיפה בהערכת חשיבה ביקורתית ויצירתית שסיכמה מחקרים מהעולם בנושא הערכת מיומנויות. סקירה זו אפשרה לבחון דרכים חדשניות להערכת מיומנויות ולגבש תשתית ידע רחבה לקראת פיתוח המבחן.

הערכה בסביבה פתוחה: נבדקו מודלים של הערכה בסביבה שאינה מנותקת מהמידע הזמין לתלמידים. זאת על סמך התנסויות בתחום בעבר בארץ ובעולם.

הערכה באמצעות אמצעי קצה שונים: נלמדו האפשרויות וההשלכות של שימוש במגוון אמצעים דיגיטליים, כגון טלפונים ניידים ומחשבים, לצורך ביצוע הערכה.

2. כינוס צוות מומחים להתייעצות

לקראת פיתוחו של המבחן גיבשה רמ"ה צוות חשיבה ראשוני, שהתכנס שלוש פעמים. הצוות הורכב מנציגים המובילים תחומי עניין מגוונים ומגופים שונים: כגון המזכירות הפדגוגית, האקדמיה ונציגות תלמידים. בחברי הצוות: **פרופ' עדו רול** (חבר סגל הפקולטה לחינוך למדע וטכנולוגיה בטכניון, וראש צוות המומחים בנושאים המיוחדים של בחינת פיז"ה ב-OECD), **ד"ר ענת בן-סימון** (לשעבר מנכ"לית המרכז הארצי לבחינות ולהערכה), **ד"ר מירב אטאלי** (מנהלת טלדור אססמנט ולשעבר הפסיכומטריקאית הראשית במט"ח), **פרופ' שרית ברזלי** (מהפקולטה לחינוך באוניברסיטת חיפה) ו**ד"ר גיורא אלכסנדרון** (מהמחלקה להוראת המדעים במכון ויצמן).

הוועדה גיבשה יחד את מאפייני המבחן ואת העקרונות המנחים לפיתוחו, ועבודתה שימשה בסיס מקצועי איתן להמשך העבודה. במאפיינים שהוגדרו בסיומו של שלב זה:

- **קהל היעד לפיתוח הראשוני:** תלמידי כיתה ט'.
- **משך המבחן:** כ-60 דקות.

- **אמצעי ההיבחנות:** העדפה למכשירי טלפון ניידים, תוך איזון עם הצרכים התפקודיים והלוגיסטיים.
 - **לוחות זמנים:** פיילוט ראשוני ומחקר חלוץ התקיימו בשנת תשפ"ה, ומבחן ארצי צפוי בתשפ"ו.
 - **משוב אפקטיבי:** המבחן יספק לבתי הספר דוח שנועד לתמוך בתהליכי ההוראה והלמידה.
- הוועדה גם גיבשה עקרונות מנחים כדי להבטיח שהמבחן יהיה רלוונטי ומועיל בעידן הנוכחי:
- **מיקוד במיומנויות:** המבחן יתמקד במיומנויות קוגניטיביות – חשיבה ביקורתית, יצירתיות, אוריינות מידע ואוריינות דיגיטלית.
 - **הערכה בסביבה פתוחה:** המבחן יאפשר שימוש בכלים זמינים, ובהם בינה מלאכותית, בדומה לנעשה בעולם האמיתי.
 - **אותנטיות והקשר:** המשימות יפותחו בהקשרים רלוונטיים לעולמם של התלמידים, תוך שילוב תחומי דעת הומניסטיים.
 - **הוגנות ושוויון:** המבחן יעוצב כך שיצמצם הטיות ויבטיח שוויון הזדמנויות לכלל התלמידים.

3. הקמת ועדת היגוי קבועה

כדי ללוות את תהליך הפיתוח של המבחן באופן מקצועי ומתמשך הוקמה ועדת היגוי קבועה. הוועדה כוללת אנשי מקצוע מובילים מתחומי האקדמיה, ההערכה והחינוך, והם מביאים עמם ידע רב ומגוון רחב של נקודות מבט. לראשות הוועדה מונתה **פרופ' ענת בן - סימון** (יושבת ראש) – לשעבר מנכ"לית מאל"ו ובעלת תואר שלישי בפסיכולוגיה. בעלת מומחיות ארוכת שנים בתחום המדידה וההערכה. עם חברי הוועדה נמנים **פרופ' ארנון הרשקוביץ** – חבר סגל בכיר בבית הספר לחינוך באוניברסיטת תל אביב, מחקריו מתמקדים בתחום היצירתיות; **ד"ר ביאטה בייגמן קליבאנוב** – מומחית לעיבוד שפה טבעית (NLP) ופיתוח כלים ממוחשבים לניתוח טקסט, עם התמחות בכתיבה באמצעות בינה מלאכותית; **יניב ירוחם מסינגר** – מנהל אגף תחומי בחירה ורב-תחומי במזכירות הפדגוגית; **איתמר נחמה וזיו בליסון** – מדריכים ארציים במזכירות הפדגוגית ורכזי התחדשות הלמידה (מיומנויות) בחטיבות הביניים; **ד"ר אלמאזה עסאמנה** – נציגת המזכירות הפדגוגית ומומחית בתחום המורשת האסלאמית; **קרן דדיה** – מונתה בעת האחרונה לחברה בוועדה, ממונה על חדשנות פדגוגית והערכה חלופית באגף לחינוך יסודי והצטרפה לוועדה על רקע עיסוקה בפיתוח מבחנים לכיתה ו'; **ורות לקס** – מורה מובילה בתחומי החדשנות ושילוב טכנולוגיות בהוראה.

לצד ועדת ההיגוי פעל צוות הפיתוח של ראמ"ה ושל חברת טלדור אססמנט: **ד"ר תמי סבג-שושן**, מנהלת תחום הערכה מעצבת באגף פיתוח כלים, ראמ"ה; **אימאן עואדיה**, מנהלת תחום מבחנים בערבית באגף פיתוח כלים, ראמ"ה; **וד"ר מירב אטאלי**, מומחית בהערכה ובפסיכומטריקה, לשעבר הפסיכומטריקאית הראשית במט"ח, בעלת ניסיון בפיתוח כלי מדידה המותאמים לחינוך ומנהלת "טלדור אססמנט".

4. ניסוי גישוש ראשוני למבחן

בחודש יוני 2024 נערך ניסוי גישוש ראשוני למבחן החדש בבית ספר בדרום הארץ. ההתנסות, שנעשתה בכיתת מחשבים נייחים, נועדה לבחון התנסות ראשונית של תלמידים במשימה שפותחה, תוך שימוש בכלי בינה מלאכותית. צוותים מורחבים מטעם טלדור אססמנט ורמ"ה ליוו את עשרת תלמידי כיתה ט' שהשתתפו במבחן.

תהליך ההתנסות: כחלק מהניסוי קיבלו התלמידים משימה בחוברת צבעונית ובה הוראות מפורטות, שהצריכה שימוש בכלי בינה מלאכותית. כל התלמידים בחרו להשתמש ב-ChatGPT, ואחת התלמידות אף השתמשה בו באנגלית באמצעות כלי תרגום. ליד כל אחד מהתלמידים ישב איש צוות שתפקידו היה לתעד את מהלך העבודה ולספק הכוונה, כדי להבין את תהליכי החשיבה שלהם בזמן אמת. התלמידים שיתפו פעולה באופן מלא וביצעו את המשימה ברצינות.

במהלך ההתנסות לא התעוררו בעיות טכניות של ממש, למעט מקרים מעטים של התנתקות מהבינה המלאכותית. התלמידים הורשו להשתמש במחשבים נייחים ובטלפונים ניידים. רובם העדיפו להשתמש במחשבים. תוצרי העבודה, שהועתקו לכונני USB, היו במסמכי Word או במצגות PPT.

תשאל ומשוב: בתשאל שנערך לאחר המבחן ציינו התלמידים כי המשימה הייתה מעניינת וכי למדו דברים חדשים על הבינה המלאכותית ועל אופן השימוש בה. תלמידה אחת, למשל, ציינה כי למדה שאי-אפשר לסמוך על הבינה המלאכותית באופן מוחלט. תלמידים שכישריהם הדיגיטליים נמוכים יותר ציינו שהמשימה הייתה מאתגרת עבורם. כלל המשתתפים העריכו כי הזמן שהוקצה למשימה (שעה וחצי) היה מספיק, אך המליצו להקצות זמן נוסף. הם ציינו שהייתה זו הפעם הראשונה שבה השתמשו בבינה מלאכותית למטרות לימודיות.

סיכום ודין: לאחר התשאל כינס צוות הפיתוח ישיבה לסיכום רשמים בשיתוף התלמידים מקרב חברי הוועדה והמורה המלווה. בישיבה עלו תובנות חשובות בנוגע לבעיות אפשריות ולשיפורים נדרשים, וזאת כדי לדייק את המשימה לפני העברתה במחקר חלוץ רחב יותר. בין השאר נדונו השאלות האם הנושא שהוצג במשימה היה קשה מדי לתלמידים, האם כדאי לחלק את המשימה לשלבים ברורים יותר והאם יש לדרוש מהתלמידים לתעד את דרך עבודתם. הממצאים שימשו בסיס לשיפור המבחן. הנתונים הוצגו לפני הוועדה, ובהתאם למשוב שהתקבל הוצע מבנה מגובש יותר למשימה.

5. בניית פלטפורמה של היבחות טכנולוגית

כדי לפתח פלטפורמה ייעודית שתתמוך בכלל דרישות המבחן נעשתה, בשיתוף מינהל טכנולוגיית מידע במשרד, התקשרות עם חברת "סקילסט" (Skillst), שפיתחה בין השאר את בית הספר הווירטואלי ואת פלטפורמות משימות הביצוע למבחני הבגרות. הפלטפורמה של סקילסט נועדה לבנייה ולניהול של תכנים לימודיים מותאמים אישית. המערכת שפותחה נועדה לבחון מיומנויות בסביבה פתוחה. הממשק של המבחן עוצב תוך מחשבה על שימוש בטלפונים ניידים כממשק העיקרי, מה שהכתיב נראות פשוטה ונקייה גם בגרסת המחשב. המטרה – לספק חוויה אחידה, אינטואיטיבית וידידותית לכלל הנבחנים.

המבחן עצמו כולל חמישה חלקים ושאלון נפרד, והוא זמין בשפות עברית וערבית. הכניסה למבחן נעשית באמצעות מערכת הזדהות אחידה, המבטיחה זיהוי וגישה מאובטחת לכל תלמיד. במהלך הבחינה המערכת דואגת לשמירת נתונים רציפה. תשובות התלמיד נשמרות מקומית ונשלחות אוטומטית לשרת 6 פעמים, בסיום כל חלק במבחן, כדי למנוע אובדן מידע במקרה של ניתוק. מנגנון זה מאפשר לתלמיד להמשיך בבחינה בדיוק מהנקודה שבה עצר. בסיום המבחן מופק פלט מפורט הכולל את תוצאותיו. פלט זה, בצירוף נתוני לוג הכוללים גם זמן שהייה בכל עמוד, מועברים באמצעות API מאובטח.

כדי להבטיח ניהול יעיל של המבחן, המערכת מבוססת על מודל תפקידים היררכי. האדמין מופקד על הניהול ברמה הארצית, כולל הגדרת זמני המבחן והזנת פרטי בתי הספר והרכזים. הרכז מטפל בניהול המבחן ברמת בית הספר, והמשגיח מנהל את המבחן בתוך הכיתה. מוקדי תמיכה ייעודיים המופעלים על ידי משרד החינוך, ראמ"ה וטלדור, דואגים לטפל בתקלות שונות בזמן אמת, החל מבעיות הזדהות ועד לתקלות בממשק. מערך תמיכה זה, בצירוף בדיקות שנעשו בסדרת פיילוט, מאפשרים לזהות ולנתח ממצאים בהתמדה כדי לשפר ולדייק את ממשק המערכת.

6. מיני-מחקר חלוץ למבחן החדש

בחודש מרץ 2025 נערך מיני-פיילוט בבית ספר במרכז הארץ, כדי לבחון את המשימה הממוחשבת באמצעי קצה שונים ולקבל תובנות ראשוניות מהתלמידים. השתתפו בו 71 תלמידי כיתה ט', שהתנסו במשימה באמצעות טלפונים אישיים ומחשבים ניידים.

המיני-פיילוט הוביל למספר ממצאים ראשוניים: המשימה הועברה בנוסח אחיד, אך נרשמו הבדלים ניכרים בין הפלטפורמות: העבודה באמצעות טלפונים אישיים הייתה שקטה יותר, ואילו העבודה באמצעות מחשבים ניידים הייתה רועשת יותר. זאת ככל הנראה עקב קשיים טכניים שתלמידים רבים חוו, כגון צילום מסך או ניווט בין חלונות. חלק מהתלמידים חוו קושי בפרטי ההזדהות האחידה של משרד החינוך, וחלקם נזקקו להכוונה בשימוש בכלים כגון בינה מלאכותית. תלמידים מעטים הגישו את המבחן מוקדם מדי, מאחר שלא הבינו שאי-אפשר לחזור אחורה לאחר המעבר בין שלבים.

בתשאל שנערך לאחר המבחן גילו תלמידים רבים עניין והתמודדו בהצלחה עם המשימות המורכבות. עם זאת המשימות היו בעיניהם "ארוכות" וחזרתיות, והם הציעו לגוון את נושאי המשימות.

לאחר ההתנסות ניתח צוות הפיתוח את התוצאות והציג אותן לוועדת ההיגוי. לנוכח הממצאים שהתקבלו במחקר חלוץ, וכדי לתת מענה לקשיים ולמשוב של התלמידים, נערכה גרסה מתוקנת של המבחן בתיאום מלא עם הוועדה. כדי להתמודד עם קשיים טכניים שופר ממשק המבחן באמצעות הוספת הנחיות ברורות יותר להזדהות ולביצוע פעולות כגון צילום מסך. כדי למנוע הגשות מוקדמות נוספה התרעה מפורשת על היעדר אפשרות לחזור לאחור ועל שלב האישור הסופי. צוות הפיתוח גם פעל לגוון את נושאי המשימות והשאלות בעקבות המשוב בדבר תחושת האורך והחזרתיות.

7. ביצוע המחקר החלוץ

בחודש יוני 2025 הועבר המבחן בפעם הראשונה במתכונת מקוונת, פתוחה ודיגיטלית בהיקף רחב. התלמידים נבחנו באמצעות מכשירים אישיים שלהם (טלפונים ניידים ומחשבים) ובאינטרנט פתוח, תוך שילוב של שימוש בכלי בינה מלאכותית. המחקר נועד לבחון את איכות השאלות, את תפקוד המערכת הטכנולוגית ואת התנהלות תפעולו של המבחן.

הדגימה והיקף המחקר: במחקר השתתפו כ-1,160 תלמידים מ-50 בתי ספר דוברי עברית שבפיקוח הממלכתי והממלכתי-דתי). תלמידים בבתי ספר דובר ערבית לא נבחנו במועד זה מפאת הפסקת לימודים על רקע ביטחוני (מבצע "עם כלביא"), ונקבע להם מועד חלופי בתחילת שנת הלימודים תשפ"ו. בחירת בתי הספר נעשתה על בסיס משתני מיון מרכזיים: סוג פיקוח (ממלכתי/ממלכתי-דתי), מחוז, רמה חברתית-כלכלית ומאפייני מגדר בבית הספר. כמו כן נכללו בתי ספר שהשתתפו בעבר במבחן "תנופה" בשפת אם או במדעים (לכיתות ט') כדי לבחון את הקשר בין המבחנים השונים. הדגימה בתוך כל בית ספר שנבחר הייתה פרופורציונלית לגודל של שכבת כיתה ט', מה שאפשר להגיע לייצוג מדויק של כל קבוצה בתוך המחקר.

תנאי הבחינה: התלמידים נבחנו באמצעות שני אמצעי קצה – **טלפון נייד אישי או מחשב**. הקצאת התלמידים לאמצעי זה או אחר הייתה אקראית.

מבנה המבחן וסוגי הפריטים: המבחן שילב בין פריטים פתוחים (עם אפשרות לציון חלקי) ופריטים סגורים (נכון/לא נכון), וכל פריט מופה בקפידה למיומנויות ולשלבי אוריינות המידע שנבדקו. בסיום המבחן ענו התלמידים על שאלות רפלקטיביות שנועדו לבחון את תחושותיהם בנוגע לנושאים שנבדקו. עם השלמת תהליך הפיתוח והאימות יפורסם המבחן באתר רמ"ה.

תהליך ההכשרה והבדיקה של פריטי המבחן: תהליך הבדיקה של הפריטים הפתוחים החל ביום הכשרה אינטנסיבי שארך שש שעות. המעריכות קיבלו רקע מקיף על מטרות המבחן ועל תהליך הבדיקה, והכירו את מבנה המבחן ואת סוגי הפריטים השונים. לאחר מכן הן התנסו בבדיקת פריטים ועשו תרגול מעשי. אחר כך, במשך כמה ימים, נערך תהליך הבדיקה בפועל. צוות המעריכות חולק לשתי קבוצות, ולכל קבוצה הוצמדה מעריכה בכירה שסיפקה הדרכה ומענה על שאלות. לאחר כל סבב בדיקה העניקה המעריכה הבכירה משוב מפורט שהבטיח עקביות ודיוק מרביים בהערכה. כמו כן, לצורך בדיקת המהימנות בין שופטים נעשתה הערכה כפולה של כ-50% מהתגובות על הפריטים על ידי שתי מעריכות בלתי תלויות.

התנהלות המבחן בפועל: רובם המכריע של התלמידים והמוסדות סיימו את המבחן ללא תקלות. להלן יובאו עיקרי התובנות והאתגרים שעלו במהלכו, וזאת כדי להצביע על נקודות לשיפור ולהפקת לקחים:

- **הזדהות וכניסה למבחן:** תהליך ההזדהות האחיד אתגר את התלמידים שלא היו רשומים או שפרטי הטלפון שלהם לא היו מעודכנים. כמו כן נרשם עומס גבוה במוקד התמיכה של משרד החינוך. עקב זאת פיתחנו מודל התחברות חלופי שיאפשר למשגיחים בכיתות להכניס תלמידים שפרטי הקשר שלהם אינם מעודכנים.

- **תאימות טכנית:** היכולת להשתמש בגמישות במגוון מכשירים ודפדפנים היא בעלת חשיבות. הנתונים העידו על צורך באופטימיזציה נוספת של המערכת, בעיקר במכשירים ישנים ובדפדפנים מסוימים, ובמתן הנחיה לתלמידים להגיע עם מכשירים טעונים. עקב זאת פיתחנו דף מוכנות ייעודי, המאפשר לתלמידים לבדוק את תאימות המכשיר והדפדפן טרם המבחן ולשדרגם לפי הצורך.
 - **יציבות במהלך המבחן:** דווח על קשיים כגון אובדן נתונים עקב ריענון הדף וקושי למצוא את כפתור ההגשה. עקב זאת נעשו שדרוגים במערכת הממוחשבת, תוך טיפול בקשיים.
 - **תשתיות בבתי הספר:** נדרש תיאום טוב יותר בין המערכת לתשתיות הפיזיות בבתי הספר, כגון זמינות מחשבים, עוצמת ה-Wi-Fi וזמינות צוותי תמיכה.
 - **כשירות המשגיחים:** תפקיד המשגיחים מצריך התאמה למתכונת הדיגיטלית. יש צורך בהכשרות ממוקדות שיעצימו אותם ויחזקו את יכולתם לסייע לתלמידים.
 - **תוכן המבחן וחווית התלמידים:** המשוב שקיבלנו על תוכן המבחן, ובכלל זה החזרתיות והאורך, הוא הזדמנות לשדרג את המשימות. יש צורך גם בהדרכה ברורה יותר בנושא השימוש בבינה המלאכותית.
 - **ייעול מוקדי התמיכה:** יש צורך בייעול ובפיתוח של מערכת מאוחדת שתגתב את הפניות ביתר יעילות ותשפר את חווית המשתמש.
- ניתוחים מתוכננים:** בעקבות המחקר החלוץ נעשו וייעשו מספר ניתוחים, ובהם:
- **ניתוח רוחבי** – יכלול בדיקות מהימנות ותוקף כללי של המבחן, השוואות בין קבוצות דמוגרפיות (סוג פיקוח, רקע חברתי-כלכלי) והשוואה בין ביצועים במחשב לבין ביצועים בטלפון נייד.
 - **ניתוח ספציפי** – יתמקד בביצועים ברמת הפריטים והמיומנויות, ויכלול ניתוח דיפרנציאלי (DIFF) כדי לזהות פריטים מוטים ובדיקת מהימנות בין שופטים עבור פריטים פתוחים.
 - **ניתוחי תוקף** – יכללו ניתוח גורמים (כדי לבדוק את מבנה המבחן), תוקף מתכנס (השוואה למבחנים בשפת אם ובמדעים) וניתוחי זמן תגובה.
 - **מודלים מתקדמים** – בחינת היתכנות של יישום מודל רב-ממדי (MIRT) ומודל דיאגנוסטי (CDM) כדי למדוד יכולות מרובות במקביל ולהפיק פרופיל יכולות אישי לכל נבחן.

תוצאות המחקר החלוץ

הניתוח הראשוני של תוצאות המחקר החלוץ מעלה ממצאים חשובים בנוגע לביצועי התלמידים ולאיכות המבחן. ככלל, במחקר השתתפו 1,160 תלמידים, שחולקו בין ארבעה נוסחים שונים של המבחן. לאחר סינון תלמידי שילוב ועולים חדשים (עד 4 שנים מתאריך העלייה) ובתי ספר שהתנדבו (ולא נדגמו), נותחו ביצועיהם של 934 תלמידים. טווח הציונים הממוצע בארבעת הנוסחים נע בין 47 ל-50 בסולם גלם (0-100 נקודות), עם סטיית תקן של 23-26.

מהימנות הפריטים וניתוחם

מהימנות פנימית: נמצא כי הפריטים במבחן בעלי מהימנות פנימית גבוהה, עם אלפא קורנברך של 0.91-0.94 (לפי הטווחים המקובלים בספרות, אלפא מעל 0.9 נחשבת גבוהה). נתונים אלו מעידים על עקביות פנימית בין הפריטים השונים, כלומר הם מודדים תכונה או מיומנות אחידה.

מתאמים בין הפריטים לציון המבחן: רוב הפריטים במבחן נמצאו בעלי מתאם גבוה מ-0.4, בעוד 4-6 פריטים נמצאו בטווח המקובל של 0.2-0.4 (פריט שהמתאם שלו עם הציון הסופי גבוה מ-0.2 נחשב בעל מהימנות סבירה). בשני פריטים נמצאה קורלציה נמוכה יותר באחד מתוך ארבעת הנוסחים (0.19 ו-0.16). הסיבה לכך שממצא זה נמצא רק באחד הנוסחים היא בלבול בין שני אתרים בעלי שם כמעט זהה ששאלות אלו התבססו עליהם. בעקבות הממצאים הוחלט לפסול את הפריטים האלה (ראו פירוט בניתוחי פריטים).

מהימנות בין שופטים: כ-50% מהתגובות על הפריטים הפתוחים נבדקו בדיקה כפולה בידי שתי מעריכות בלתי תלויות, והוצג בהם אחוז הסכמה ממוצע של 89.6%, מתאם ממוצע של 0.85 וקאפה ממוצעת של 90.7. כל אלה מצביעים על רמת הסכמה גבוהה בין המעריכות. הנתונים מעידים כי ברוב השאלות הקריטריונים של הניקוד היו ברורים ואפשרו בדיקה עקבית ואמינה. בשלושת הפריטים שנמצא כי מהימנותם נמוכה יותר עשינו שיפורי מחוון כך ששיטת הציון תהיה דומה יותר לפריטים שבהם המהימנות הייתה גבוהה.

ניתוחי פריטים: בבדיקות הפריטים נמצא שבשני פריטים יש בעיה בניקוד החלקי של השאלה. בשתי השאלות היו במקור 4 קטגוריות ציון (ובהן שתי קטגוריות של ניקוד חלקי שמייצגות תשובה חלקית על השאלה). בהסתכלות על שיעור התלמידים שקיבלו את הניקוד החלקי ועל הממוצעים שלהם בציון המבחן עצמו, נמצא כי מיעוט מקרב התלמידים קיבלו את הניקוד החלקי, וכי הממוצעים שלהם דומים מאוד. לכן הוחלט לאחד את קטגוריות הציון החלקי, כך שבמקום 4 קטגוריות ציון יהיו רק 3 קטגוריות ציון (איחוד הקטגוריות האמצעיות). כמו כן הוחלט לפסול שני פריטים עקב מהימנות נמוכה באחד הנוסחים. התפקוד הבעייתי בפריטים אלו נבע מכך שהתלמידים התייחסו לאתר "עבדת" במקום ל"עין עבדת". לאחר יישום ההחלטות בעקבות ניתוחי הפריטים חושבו הציונים מחדש ונעשה כיוול קלאסי בין הנוסחים.

השוואות בין נוסחים ואוכלוסיות

השוואת נוסחים: נעשו בדיקות DIF (ניתוח תפקוד פריט דיפרנציאלי) בין ארבעת נוסחי המבחן כדי לוודא שאין הבדלים מפלים ביניהם. הבדיקות הראו התאמה בין רוב הפריטים, למעט שני פריטים בנוסח השני שזוהו כיוצאי דופן והוסרו (ראו פירוט בניתוחי פריטים).

הבדלי מגדר: לא נמצא אפקט מובהק למגדר, $t = 0.49, p = 0.623, \beta = 0.01$ נראה כי בשני פריטים הבנות השיגו ביצועים טובים במעט מהבנים (הפרש של 0.1-0.12 באחוז ההצלחה בפריטים אלו בקרב הבנות).

הבדלים בין סוגי פיקוח: נמצא אפקט מובהק לפיקוח לטובת תלמידי החינוך הממלכתי בהשוואה לתלמידי החינוך הממלכתי-דתי, $t = -2.7, p = 0.007, \beta = -0.09$

הבדלים על בסיס רקע חברתי-כלכלי: נמצא הבדל מובהק לרקע חברתי-כלכלי. תלמידים מרקע חברתי-כלכלי גבוה (קטגוריה 1) הצליחו יותר במבחן מתלמידים מרקע בינוני (קטגוריה 2), ותלמידים מרקע חברתי-כלכלי בינוני הצליחו יותר מתלמידים מרקע נמוך (קטגוריה 3), $t = -5.5, p < 0.001, \beta = -0.17$

להלן ממוצעי המבחן לפי פילוחים (הממוצעים וסטיות התקן של הציונים הם עבור ציוני גלם בסולם 0-100 לאחר כיוול בין נוסחים ויישום החלטות ניתוחי פריטים, והם ממושקלים באמצעות משקולות הדגימה):

test score sd	test score mean	
24.27	51.88	כללי
25.51	50.95	בנים
23.08	52.71	בנות
23.86	53.64	פיקוח ממלכתי
24.68	45.92	פיקוח ממלכתי-דתי
23.60	56.58	רקע חברתי-כלכלי גבוה
24.27	51.06	רקע חברתי-כלכלי בינוני
23.74	45.39	רקע חברתי-כלכלי נמוך

השפעה של ידע קודם על הצלחה במבחן

לצורך הערכת התרומה של התנסות קודמת ביישומים דיגיטליים ושל היכרות מעמיקה עם עולם התוכן של המבחן (תכנון טיולים) על הציון הכולל, התלמידים מילאו שאלון קצר הבודק את מידת ההיכרות הקודמת שלהם עם יישומים דיגיטליים (שימוש בבינה מלאכותית, מנועי חיפוש, רשתות חברתיות וכו'), ושאלון קצר אחר שבדק את מידת הרלוונטיות של עולם התוכן עבורם (ניסיון

בתכנון טיולים, סקרנות ועניין בנושא טיולים, באיזו מידה התלמיד נוהג לטייל). נבדקה רגרסיה מרובה שבה הציון במבחן מנובא באמצעות משתני הרקע של התלמיד (מגדר, סוג פיקוח, רקע חברתי-כלכלי) והדיווח העצמי של התלמידים על מידת הידע הקודם שלהם ביישומים דיגיטליים ועל מידת ההיכרות שלהם עם עולם התוכן.

השפעה של התנסות קודמת ביישומים דיגיטליים על הצלחה במבחן: נמצא כי מידת ההיכרות הקודמת עם יישומים דיגיטליים היא מנבא מובהק, $t = 3.18, p = 0.001, \beta = 0.11$ ממצא זה הוא עדות מסוימת לתוקף המבחן ככלי שמודד אוריינות דיגיטלית.

השפעה של היכרות קודמת עם עולם התוכן של המבחן על הצלחה המבחן: נמצא כי היכרות קודמת עם תכנון טיולים היא מנבא מובהק, $t = 2.89, p = 0.003, \beta = 0.10$ נראה כי עולם התוכן שבו עוסק המבחן עשוי להשפיע על מידת הצלחתם של התלמידים. נראה אפוא שזה אתגר שנצטרך להתמודד איתו בפיתוחים בעתיד: יהיה עלינו להקפיד על שימוש בעולם תוכן רחב או ניטרלי ככל האפשר, או לחלופין להשתמש במספר עולמות תוכן שונים במבחן אחד, כך שציון המבחן לא יושפע באופן בלעדי ממידת ההיכרות המוקדמת עם עולם תוכן זה או אחר.

הבדל בין תנאי היבחנות

נבדקה ההשפעה של אמצעי ההיבחנות (טלפון נייד לעומת מחשב) על ההצלחה במבחן. התלמידים סווגו באקראי לשני תנאי היבחנות: מחציתם ענו על שאלות המבחן באמצעות מחשב, ומחציתם באמצעות הטלפון האישי שלהם. נמצא הבדל של כ-0.38 נקודות בציון בין התנאים, הבדל שאינו מובהק, $t(1, 773) = 0.23, p = 0.81$.

תוקף מתכנס ותוקף מבחן

מתוך 934 התלמידים שהשתתפו במחקר – 424 נבחנו גם במבחן תנופה בשפת אם לכיתות ט', ו-379 נבחנו גם במבחן תנופה במדעים לכיתות ט', שהועברו בתשפ"ה. כדי לבדוק את התוקף המתכנס והמבחן של המבחן נבדקו המתאמים בין ציוני המבחן לציונים במבחני תנופה ובתת-הנושאים (בשפת אם: הבנת הנקרא, לשון, חשיבה יצירתית-ביקורתית. במדעים: ביולוגיה, כימיה, אנרגיה, חשיבה ביקורתית-יצירתית).

המתאם של הציון במבחן עם מבחני תנופה בינוני בגודלו ($r = 0.49$ בשפת אם, $r = 0.52$ במדעים), באופן יחסי דומה למתאם בין מבחנים בנושאים שונים (נמוך ממנו במעט)¹. המתאם של ציון המבחן עם תת-נושא חשיבה ביקורתית-יצירתית גם הוא בינוני בגודלו ($r = 0.48$ בשפת אם, $r = 0.46$ במדעים), בין השאר מכיוון שיש מתאם גבוה בין ציוני המבחנים לציון בתת-נושא חשיבה ביקורתית-יצירתית ($r = 0.71$ בשפת אם, $r = 0.83$ במדעים). על כן בדקנו את המתאמים בציונים שעבורם אין חפיפה בין פריטים. למשל, במבחן בשפת אם אין חפיפה בין פריטי תת-נושא לשון ותת-נושא חשיבה ביקורתית-יצירתית, ומצאנו כי המתאם של ציון המבחן עם תת-נושא לשון הוא $r = 0.39$ בעוד המתאם של ציון המבחן עם חשיבה ביקורתית-יצירתית הוא $r = 0.48$. נוסף על כך, במבחן במדעים אין חפיפה בין פריטי ידע לבין תת-נושא חשיבה ביקורתית, ומצאנו כי המתאם של

¹ לדוגמה, במבחני המיצ"ב 2018 לכיתות ח', המתאמים בין הציונים בשפת אם, אנגלית, מתמטיקה ומדעים היו בטווח של 0.55 - 0.73.

ציון המבחן עם פריטי הידע הוא $r=0.23$, בעוד המתאם של ציון המבחן עם חשיבה ביקורתית הוא $r=0.42$.

ממצאים אלו הם עדות מסוימת לתוקף מתכנס (מדידה של חשיבה ביקורתית-יצירתית במבחני תנופה) ותוקף מבחין (מדידה של פריטי לשון במבחן תנופה בשפת אם ושל פריטי ידע במבחן תנופה במדעים).

מקורות והפניות

משרד החינוך. (2023, 3 ביולי). חוזר מנכ"ל, הוראת קבע מס' 0398: מערך ההערכה והמדידה במערכת החינוך לשנים תשפ"ד-תשפ"ו .

<https://apps.education.gov.il/Mankal/Horaa.aspx?siduri=499>

משרד החינוך. (2024, 1 בספטמבר). חוזר מנכ"ל, הוראת קבע מס' 0424: מערך ההערכה והמדידה במערכת החינוך .

<https://apps.education.gov.il/Mankal/horaa.aspx?siduri=526>

משרד החינוך. (2024, 1 בדצמבר). חוזר מנכ"ל, הוראת קבע מס' 0429: מערך ההערכה והמדידה במערכת החינוך .

<https://apps.education.gov.il/Mankal/horaa.aspx?siduri=532>

משרד החינוך. (תשפ"א-תשצ"א). תפיסת הלמידה המתחדשת במערכת החינוך. ירושלים: משרד החינוך.

<https://meyda.education.gov.il/files/Planning/dmuthabogermismach.pdf>

משרד החינוך. (תשפ"א-תשצ"א). תפיסת הלמידה המתחדשת במערכת החינוך – מיומנויות. ירושלים: משרד החינוך

<https://meyda.education.gov.il/files/Planning/dmuthabogeravneiderech.pdf>

Ananiadou, K., & Claro, M. (2009). *21st Century Skills and Competences for New Millennium Learners in OECD Countries* (OECD Education Working Papers, No. 41). OECD Publishing. <https://doi.org/10.1787/218525261154>

Greiff, S., Holt, D. V., & Funke, J. (2014). Perspectives on problem solving in educational assessment: Analytical, interactive, and collaborative problem solving. *Journal of Problem Solving*, 7(1), 71–91.

<https://docs.lib.purdue.edu/cgi/viewcontent.cgi?article=1153&context=jps>

Leaning, M. (2019). An approach to digital literacy through the integration of media and information literacy. *Media and Communication*, 7(2), 4-13.

https://www.researchgate.net/publication/333721730_An_Approach_to_Digital_Literacy_through_the_Integration_of_Media_and_Information_Literacy

Lodge, J. M., Howard, S., & Broadbent, J. (2023). [Assessment redesign for generative AI: A taxonomy of options and their viability.](#)

Schäfer, J., Reuter, T., Karbach, J., & Leuchter, M. (2024). Domain-specific knowledge and domain-general abilities in children's science problem-solving. *British Journal of Educational Psychology*, 94(2), 346-366.

<https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bjep.12649>

Sternberg, R. J. (2021). Adaptive intelligence: Its nature and implications for education. *Education Sciences*, 11(12), 823.

<https://www.mdpi.com/2227-7102/11/12/823>

Weldon, P. R. (2019). *Changing priorities? The role of general capabilities in the curriculum*. Australian Council for Educational Research (ACER).

<https://research.acer.edu.au/cgi/viewcontent.cgi?article=1002&context=nswcurriculumreview>

OECD (2017). *PISA 2015 Results (Volume V): Collaborative Problem Solving, PISA, OECD Publishing, Paris*.

<https://doi.org/10.1787/9789264285521-en>.

OECD (2019). *OECD Future of Education and Skills 2030 Conceptual learning framework*, OECD Publishing, Paris

אנשי הצוות

מנהלת המהלך: ד"ר תמי סבג-שושן, מנהלת תחום הערכה מעצבת באגף פיתוח כלים, ראמ"ה

ליווי תהליך הפיתוח: אימאן עואדיה, מנהלת תחום מבחנים בערבית באגף פיתוח כלים, ראמ"ה

יו"ר ועדת ההיגוי: ד"ר ענת בן-סימון (לשעבר מנכ"לית מאל"ו)

חברי ועדת ההיגוי: יניב מסנגר, איתמר נחמה, ד"ר אלמאזה עסאמנה וקרן דדיה (המזכירות הפדגוגית), פרופ' ארנון הרשקוביץ (ראש החוג לחינוך מתמטי, מדעי וטכנולוגי בבית הספר לחינוך באוניברסיטת תל אביב), ד"ר ביאטה בייגמן קלבנוב (Principal Research Scientist at Educational Testing Service – ETS), רות לקס (מורה).

צוות פיתוח המבחן: ד"ר מירב אטאלי, נדב כספי, לירון גביסון, מילה ברוסין (טלדור אססמנט).

עיבוד תוצאות המבחן: נורית ליפשטט, ד"ר לימור בורוט וד"ר פרנקלין זרומב (ראמ"ה), ד"ר צור קרליץ, נעם כאהן, איתמר כהן, נתנאל סילברשטיין ורועי אקסלרוד (מאל"ו).

תהליך הטמעת המבחן: ליאת ארנהיים, מנהלת אגף השפעה, ודנה אלון הנדלר, מנהלת תחום קהילות ושותפויות, ראמ"ה.

מנכ"ל ראמ"ה: ד"ר גל אלון