

מחקר פיזה 2025 (PISA)

דוח טכני: מאפיינים, כלי הערכה ומהלך המחקר



ספטמבר 2026, אלול תשפ"ו



משרד החינוך

תוכן העניינים

1. מבוא	1.1 מהו הדוח הטכני ומה מקומו?	3	1.2 רקע כללי - מהו מחקר פיזה?	4
2. מאפיינים ייחודיים של מחקר פיזה 2025	2.1 אוריינות מדעית כתחום מרכזי	5	2.2 למידה בעולם הדיגיטלי	5
	2.3 הערכת אנגלית כשפה זרה	6	2.4 ישראל בתקופת מלחמה	7
3. מסגרת הערכה והמודל הקוגניטיבי במדעים	3.1 הגדרת אוריינות מדעית בפיזה 2025	8	3.2 מה חדש במסגרת המושגית בתחום המדעים של פיזה 2025?	8
	3.3 המודל הקוגניטיבי - שלוש הכשירויות המדעיות	9	3.4 סוגי הידע הנדרשים באוריינות מדעית	10
	3.5 זהות מדעית, עמדות ופעלונות סביבתית	11	3.6 הקשרי המשימות	12
4. תיאור הערכה וסוגי הפריטים	4.1 מבנה יחידות המבחן	13	4.2 סוגי הפריטים	13
5. עיצוב ופיתוח ההערכה ומחקרי התיקוף	5.1 תרגום והתאמה לישראל	14	5.2 מחקרי תיקוף על ידי OECD	14
	5.3 מחקר החלוץ בישראל	14	5.4 הכשרת הבודקים	15
6. הליכי הבדיקה והקידוד	6.1 תהליכי קידוד	16	6.2 מהימנות בין בודקים	16
7. מערך המחקר בישראל: דגימה, העברה ובקרת איכות	7.1 אוכלוסיית היעד בישראל	17	7.2 ההשפעה של המצב הביטחוני על ביצוע המחקר	17
	7.3 הליך הדגימה הדו-שלבי	17	7.4 גריעה מאוכלוסיית היעד	17
	7.5 שכבות הדגימה בישראל	17	7.6 שיטת הדגימה	18
	7.7 מספר המשתתפים ושיעורי ההשתתפות	18	7.8 השתתפות בתי ספר בישראל והשתתפות תלמידים	19
	7.9 שפות ההערכה בישראל	20	7.10 הליך ההיבחנות ואיסוף הנתונים	20
	7.11 ניהול הנתונים, אבטחה ושמירה	20	7.12 בקרת איכות על ביצוע המבחן	20
8. שאלוני הרקע, מערך המדידה והניתוח הסטטיסטי	8.1 מדד חברתי-תרבותי-כלכלי של פיזה	21	8.2 מערך המדידה, הכיול והניתוח הפסיכומטרי	21
9. רמות הבקיאות	9.1 רמות הבקיאות באוריינות מדעית	24	9.2 רמות הבקיאות באוריינות קריאה	25
	9.3 רמות הבקיאות באוריינות מתמטיקה	26		

מבוא

1

מהו הדוח הטכני ומה מקומו?

1.1

הדוח הטכני של מחקר פיזה 2025 מתפרסם לצד דוח ממצאי המחקר. שני הדוחות משלימים זה את זה. דוח הממצאים הלאומי מציג את תוצאות ישראל במחקר פיזה 2025: ציוני התלמידים, מיקומה של ישראל ביחס למדינות אחרות, פערים בין קבוצות אוכלוסייה, שינויים לאורך זמן ומסקנות מדיניות עיקריות. דוח זה מיועד לקובעי מדיניות, לאנשי חינוך, לחוקרים ולציבור הרחב.

לצידו, הדוח הטכני מציג את "מאחורי הקלעים" של התהליך המחקרי והרציונל המקצועי: כיצד תוכנן וגובש המחקר, מה נמדד ומדוע, כיצד פותחו הפריטים, כיצד בוצעה הדגימה, כיצד קודדו התשובות ומהן רמות הבקיאיות שמאפשרות לפרש את הציונים. הדוח הטכני אינו עוסק בתוצאות עצמן, אלא בתשתית המדעית והמתודולוגית שעליה הן נשענות.

לדוח הטכני שתי מטרות:



חשוב לציין כי המסגרת המושגית, כלי ההערכה, שיטות הקידוד ונוהלי הדגימה פותחו על ידי מינהל החינוך של ה-OECD כחלק מתהליך שבו לקחו חלק המדינות המשתתפות במחקר. ראמ"ה, כמרכז המקצועי הלאומי של ישראל לנושאי המדידה וההערכה של מערכת החינוך, חברה במועצה הבינלאומית של המחקר (PISA Governing Board); לוקחת חלק במפגשי הפיתוח, התפעול והעיבוד השונים; שותפה בפורום מנהלי המחקר (National Project Managers); ופועלת בהתאם לנוהלי המחקר (PISA Technical Standards). תחת הנהלים הבינלאומיים, ראמ"ה מתאימה את המחקר ככל הניתן לצרכים המקומיים הייחודיים ואחריות על יישומו בישראל.

1.2 רקע כללי – מהו מחקר פיזה?

מחקר PISA - Programme for International Student Assessment, ובעברית להלן "פיזה", נערך החל משנת 2000 על ידי מינהל החינוך של ה- OECD (Organisation for Economic Co-operation and Development). המחקר נערך במחזוריות של אחת לשלוש או לארבע שנים. זהו המחקר הבינלאומי המקיף ביותר בתחום החינוך מבחינת מספר המדינות המשתתפות בו והיקף התחומים הנמדדים בו והוא נחשב למוביל בתחום מחקרי החינוך.

במחזור האחרון (2025) השתתפו במחקר מעל 760,000 תלמידים מ-91 מדינות וישויות כלכליות המייצגים כ-33 מיליון בני ובנות 15 במקומות אלו. המבחן נערך באופן ממוחשב וארך כשעתיים כאשר כל תלמיד קיבל מקבץ שונה של שאלות מתוך המאגר הכללי. בתום המבחן ענו התלמידים על שאלון עמדות שארך 35 דקות, ועסק במגוון נושאים, בהם דעותיהם ותפיסותיהם ביחס לסוגיות שונות, הרקע ממנו הגיעו וחוויות הלמידה שלהם. גם המנהלים של אותם בתי הספר מילאו שאלון שעסק בהיבטים של ניהול וארגון סביבות הלמידה. מטרת המחקר היא להעריך את רמת האוריינות של תלמידים ותלמידות בני 15 בכל מדינה בהתבסס על מסגרת מושגית משותפת ומוסכמת המאפשרת השוואה ברמה הבינלאומית. במחקר מעריכים את יכולתם של התלמידים והתלמידות להשתמש בידע, בכישורים ובמיומנויות באופן המאפשר להם להתמודד בצורה מיטבית וביעילות עם סביבתם ולפתור בעיות בחיי היום יום. תפיסה זו מציגה השקפת עולם שונה מזו שהייתה נהוגה בעבר במבחנים סטנדרטים רחבי היקף במערכות חינוך רבות בעולם כולל בישראל. אם בעבר מטרת המבחנים הייתה להעריך את מידת השליטה של התלמידים בידע שלמדו ושינו כפי שהוא מוגדר בתוכניות הלימודים, מבחני פיזה חותרים להעריך תוצר חינוכי מורכב ורחב הרבה יותר, כשבבסיס גישה זו ההכנה לחיים במובנם הרחב: הקנייתם של כלי למידה וחשיבה, לצד היכולת ליישם ידע, מיומנויות וכישורים בתיזמון ובהקשר מתאימים. במסגרת המבחן נבדקים שלושה תחומי אוריינות: קריאה, מתמטיקה ומדעים, ובכל מחזור מחקר מושם דגש על אחד מהם כנושא המרכזי. במחזור מחקר 2025 התחום שהושם עליו דגש הוא אוריינות מדעית. כמו כן, לראשונה נוספו גם תחום האנגלית כשפה זרה, שעד כה לא היה חלק ממבחן פיזה, ותחום אורח חדש שפותח "למידה בעולם הדיגיטלי" (Learning in the Digital World, LDW).

ישראל השתתפה לראשונה במחקר פיזה בשנת 2002, עת התקיים הסבב השני של מחזור פיזה הראשון משנת 2000. במחזור המבחן שהתקיים בשנת 2003 ישראל לא השתתפה, ומאז 2006 ישראל משתתפת באופן סדיר בכל מחזורי המבחן, וזה הנוכחי הוא השמיני במספר.

כאמור מלבד הערכת רמות האוריינות נאסף במחקר מידע משלים על מאפיינים שונים של התלמיד, של משפחתו ושל המוסד החינוכי שבו הוא לומד, וזאת באמצעות שאלונים לתלמידים ולמנהלי בתי הספר. איסוף משתני רקע מסייע בהסבר השונות בהישגים הלימודיים בין תלמידים על בסיס מאפייני רקע כגון מגדר, רקע חברתי-תרבותי-כלכלי, אסטרטגיות למידה, פרקטיקות הוראה ומאפיינים הקשורים לבית הספר ולסביבה הלימודית.

גיל 15 נבחר על ידי מנהלי מחקר פיזה הואיל ובגיל זה מסתיים חינוך החובה במרבית מדינות ה-OECD, כמו גם בשל התובנה שהתגבשה לאורך מחקרים שונים, אשר לפיה מגיל זה ניתן כבר להסיק על רמת הכשירויות האורייניות שרכשו התלמידים, ושאותן יישאו עימם בהמשך חייהם כאזרחים בוגרים. הגיל המדויק של התלמידים המשתתפים בפיזה הוא מ-15 שנה ושלושה חודשים ועד 16 שנה וחודשיים.

שלושת תחומי הליבה של מבחן פיזה



מאפיינים ייחודיים של מחקר פיזה 2025

2

אוריינות מדעית כתחום מרכזי

2.1

המדע הוא אחד הכלים המרכזיים שבאמצעותם בני אדם מבינים את העולם, מפרשים תופעות, מעריכים ראיות ומקבלים החלטות מושכלות. בעידן הנוכחי, המתאפיין בהצפת מידע ובחלקו אף כוזב, הצורך באוריינות מדעית הוא קריטי, שכן היא מאפשרת פיתוח חשיבה ביקורתית ומבוססת ראיות. בנוסף, אנו חיים בתקופה שבה לסוגיות מדעיות וטכנולוגיות השפעה מכרעת על חיי היומיום של קהילות ומדינות ברחבי העולם. השפעה זו ניכרת בתחומי הבריאות, משאבי הטבע, הסביבה ושינויי האקלים, סיכונים שונים כמו לדוגמה בתחומי האנרגיה, הביטחון, המזון, ובכל הנוגע לחדשנות ולאתגרים מדעיים וטכנולוגיים בעת הזאת. על הרקע הזה ברור כי הכשרת אזרחים בעלי אוריינות מדעית היא ערך חינוכי מרכזי. במבחן 2025 עומדת אוריינות המדעים כתחום המרכזי של ההערכה, בדומה למחזורי 2006 ו-2015. המשמעות היא שלתחום המדעים מוקצה חלק ניכר מזמן המבחן ומפותחים עבורו פריטים חדשים רבים יותר, דבר המאפשר לבחון ולחקור תת-נושאים במדעים. מסגרת ההערכה של המדעים עודכנה לקראת מחזור מבחן זה, כפי שיפורט בהמשך.

למידה בעולם דיגיטלי

2.2

מחקר פיזה 2025 מציג לראשונה את תחום "למידה בעולם דיגיטלי" (LDW - Learning in the Digital World) כתחום הערכה אורח למחזור 2025. תחום זה מוגדר **כמדידת יכולתם של תלמידים בני 15 לנהל תהליך עצמאי, מחזורי ובעל מכוונות עצמית של הבניית ידע ופתרון בעיות באמצעות כלים חישוביים**. בשונה מכלי הערכה מסורתיים הבוחנים שיחזור של ידע קיים, כלי זה מתמקד באסטרטגיות חקר והבנה של מודלים דיגיטליים. הוא מעריך את הדרך שבה הלומד רוכש ומיישם הבנה מושגית חדשה כדי לפתור בעיות מורכבות ופתוחות בתוך מרחבים ממוחשבים אינטראקטיביים המדמים תרחישים מציאותיים. בבסיס המחקר בתחום זה שתי מטרות:

להבין האם תלמידים מסוגלים לפתור בעיות באמצעות כלים ופרקטיקות חישוביות בסביבות למידה דיגיטליות; לספק ראיות בנות-השוואה על השימוש בטכנולוגיות דיגיטליות בבית הספר, ולקשור בין יישומן של טכנולוגיות אלו לבין תוצרי למידה בתחומי הליבה שהוגדרו בפיזה ובתחומי למידה מבוססי-מחשב. הערכת התחום מבוססת על מודל הכשירות המבוסס על שני מרכיבים - פרקטיקות חקר ופתרון בעיות, ותהליכי למידה עצמיים:

פרקטיקות של חקר ופתרון בעיות חישוביות



ממד זה בוחן את יכולתם של התלמידים לתכנן ולבצע ניסויים מבוקרים בסימולציה, לנתח נתונים וייצוגים גרפיים משתנים כדי לזהות דפוסים וקשרים, ולבנות או למפות מודלים דיגיטליים המייצגים את פתרון הבעיה.

תהליכי למידה עצמיים



ממד זה מתמקד ביכולות המטא-קוגניטיביות, ההתנהגותיות והרגשיות של הלומד בסביבה פתוחה. הוא מעריך אסטרטגיות של ניטור התקדמות והתאמת פעולה (זיהוי פערי ידע, שימוש יעיל במשוב עצמי או בעזרי למידה חיצוניים כגון מדריכים ודוגמאות), לצד ניהול רגשי והתנהגותי (התמדה והשקעת מאמץ מול קושי או תסכול) והערכה עצמית רפלקטיבית בסיום המשימה.

הערכת אנגלית כשפה זרה

2.3

רציונל ומטרות כלי ההערכה

במחזור מחקר פיזה 2025 שולב לראשונה מבחן הרשות להערכת מיומנות השפה האנגלית כשפה זרה (FLA - Foreign Language Assessment). בעידן הכלכלה הגלובלית והסביבה הדיגיטלית המשתנה בקצב מהיר, היכולת לתקשר, לצרוך מידע ולפעול בשפה האנגלית היא נדבך חיוני במוכנותם של אזרחים צעירים לחיים הבוגרים ולשוק העבודה העתידי. הערכה זו מוגדרת כיכולתו של התלמיד/ה להשתמש בשפה זרה כדי לתקשר בצורה יעילה במצבי חיים אמיתיים. ארגון ה-OECD רואה במיומנות זו כלי חיוני המשמש שער לניידות חברתית וכלכלית והמעניק נגישות לידע אקדמי וטכנולוגי בינלאומי. מיומנות שפה ברמה סבירה דורשת שילוב דינמי בין מיומנויות קליטה באמצעות קריאה והאזנה לבין מיומנויות הפקה אקטיביות, ובעיקר יכולת להתבטא ולדבר בצורה מובנת והולמת. הרציונל המקצועי של המבחן מבוסס על מדדים בינלאומיים מובילים (כמו מדרג ה-CEFR), אך מתאים את המשימות באופן ייחודי לפרופיל תחומי העניין ולבשלות הקוגניטיבית של בני ה-15.

"המבחן מעריך את יכולת ההבעה בעל-פה ואת הבנת הנקרא של טקסטים ברמות מורכבות שונות ואת הבנת הנשמע מתוך קטעי אודיו שבהם דוברים במגוון מבטאים בינלאומיים, וזאת באמצעות ממשק דיגיטלי המאפשר להציג משימות דינמיות וגירויים אינטראקטיביים. המשימות משובצות בהקשרים רלוונטיים המשקפים את חייהם של בני הנוער, כגון ניהול אינטראקציות חברתיות, צריכת מדיה ותוכן דיגיטלי ושימוש באנגלית לצורך פרויקטים לימודיים. כאמור, בניגוד למבחני השפה הפנימיים של מערכת החינוך בישראל, מבחן ה-FLA אינו בוחן עמידה בתוכנית לימודים מסוימת, אלא מודד את רמת הבקיאות המעשית בשפה על פי אמות מידה אחידות המשוות בין מדינות שונות.

הגדרת אוכלוסיית היעד והרחבת המדגם

אוכלוסיית היעד של מבחן ה-FLA נגזרה כתת-מדגם מתוך קבוצת התלמידים שנבחרו להשתתף במחקר פיזה 2025. עם זאת, כדי שתלמיד שנדגם במדגם הלאומי יוכל להידגם גם למבחן באנגלית, עליו ללמוד אנגלית במסגרת פורמלית בתוך מערכת החינוך במשך תקופה מצטברת של שנת לימודים אקדמית אחת לפחות טרם מועד הבחינה. כתוצאה מההשתתפות בחלק זה של המחקר בוצעו שתי התאמות:

הגדלת המדגם הארצי: בעוד המחקר בשלושת תחומי הליבה דרש השתתפותם של לפחות 6,300 תלמידים ברמה הלאומית, הוספת רכיב ה-FLA חייבה להגדיל את המדגם ב-1,650 תלמידים נוספים לפחות. בכך, נדרש המדגם הישראלי לכלול לפחות 7,950 תלמידים.

שיעורי השתתפות: כדי להבטיח את התוקף הסטטיסטי של הנתונים ולמנוע הטיות השבה, נקבע כי שיעור ההשתתפות המשוקלל לתלמידים גם במחקר ה-FLA חייב לעמוד על 80% לפחות. כמו כן, בית ספר נכלל כמשתתף במבחן באנגלית, אך ורק אם שיעור ההשתתפות הלא-משוקלל של התלמידים שבו הזכאים להיבחן עלה על 33%. בתי ספר שלא עמדו בתנאי זה נגרעו ממסד הנתונים של המחקר כולו. כבכל מבחן, ובשל אופיו כמעריך מיומנות שפה זרה, הגדירה מנהלת פיזה הבינלאומית מנגנוני הגנה מתודולוגיים קשיחים שיושמו בקפדנות על ידי ראמ"ה בשלב ההיבחנות בבתי הספר:

ניטרול השפעת המורה: כדי להבטיח את אמינות תנאי הבחינה ומניעת סיוע ישיר או עקיף, חל איסור מוחלט על מורים לאנגלית או מורים המלמדים תחומי דעת מגוונים בשפה האנגלית לשמש כמשגיחים, כבוחנים או כנאמני בדיקה במוסד החינוכי שבו הם מלמדים, ובפרט בכיתות שבהן נבחנו תלמידיהם הישירים.

מניעת הכנה ממוקדת: בהתאם להנחיות הארגון החלים על כל מרכיביו של המבחן, גם בתחום האנגלית נאסר על בתי הספר לקיים שיעורי הכנה ממוקדים, מרתונים או תרגולים המבוססים על פריטי אנגלית הדומים למבנה ה-FLA. בתי הספר הורשו להשתמש אך ורק בחומרי ההסברה הכלליים וביחידות ממחקר החלוץ שנערך שנה לפני כן ופורסמו על ידי ה-OECD.

ישראל בתקופת מלחמה

2.4

המחקר בישראל נערך בתקופה חריגה במיוחד מבחינה ביטחונית, חברתית וחינוכית. מאז 7 באוקטובר 2023 פעלה מערכת החינוך הישראלית בצל מלחמת "חרבות ברזל", שבמסגרתה פונו יישובים רבים באזור הצפון והדרום למשך חודשים רבים. החיים במדינה כולה מאופיינים בחוסר יציבות מתמשך ואי-עקביות בלמידה, עבודה מרחוק, מחסור בכוח אדם הוראתי ושינויים מהותיים בשגרת החיים של תלמידים רבים. נסיבות אלה לא השפיעו על מבנה המבחן, רמתו או היקפו, אולם הן רקע להתייחסות אל הנתונים בישראל ואל התמונה המצטיירת מן ההשוואה הבינלאומית. סביר מאוד להניח שלאירועים אלה הייתה השפעה על ביצועי התלמידים, אך חשוב לזכור שאי-אפשר למדוד ישירות את השלכותיהם על ההישגים, מה גם שהמחקר הבינלאומי אינו מאפשר לקבוע קשרים סיבתיים בכל הנוגע לשינויים בהישגים.

מסגרת ההערכה והמודל הקוגניטיבי במדעים

3

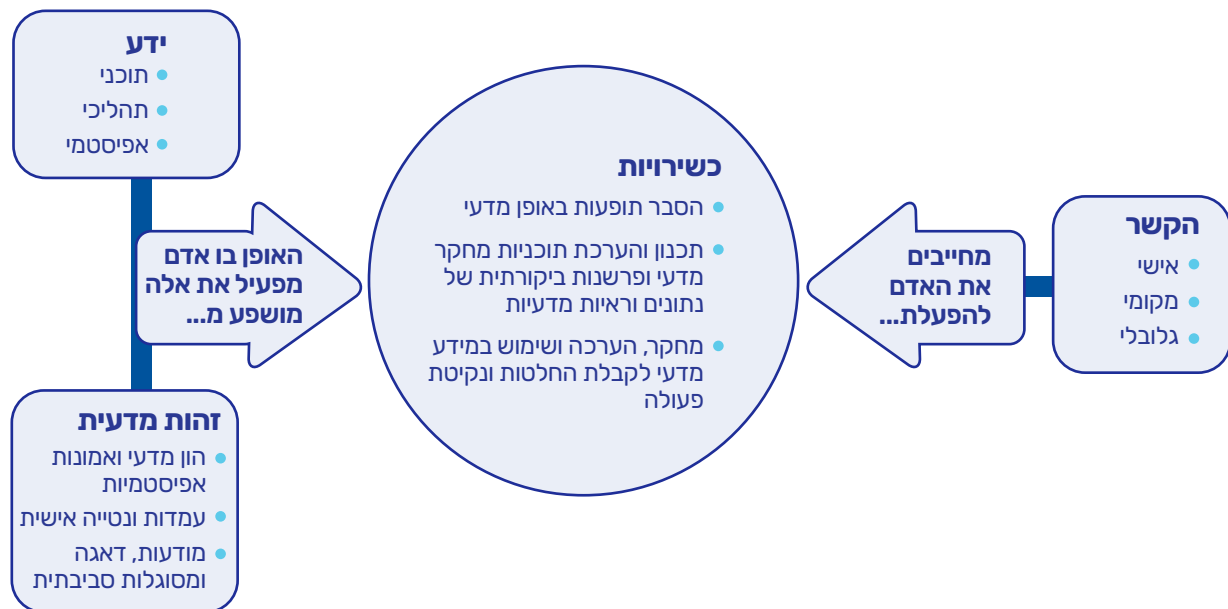
הגדרת אוריינות מדעית בפיזה 2025



יכולת לעסוק בסוגיות מדעיות ובדיון מבוסס ראיות כאזרחים חושבים; להסביר תופעות באופן מדעי; לבנות, להעריך ולפרש שאלות ותכנוני חקר מדעי; נתונים וראיות; ולאתר, להעריך ולהשתמש במידע מדעי לצורך קבלת החלטות ופעולה.

3.1

הגדרה זו משקפת תפיסה רחבה של אוריינות מדעית הכוללת ארבעה רכיבים הקשורים זה בזה: (1) ההקשרים שבהם מוצגות המשימות; (2) סוגי הידע המדעי שהתלמיד נדרש להפעיל; (3) הכשירויות המדעיות הנמדדות; ו-(4) עמדות וזהות מדעית. שלושת הרכיבים הראשונים נמדדים באמצעות פריטי המבחן הקוגניטיבי עצמו, ואילו הרביעי נמדד בעיקר באמצעות שאלוני רקע ועמדות.



מטרת ההערכה במדעים מתמקדת ביכולת ליישם ידע מדעי בהקשרים חדשים ומגוונים, ולא רק בשחזור הידע לבדו או את השליטה של התלמידים בעובדות, במושגים או בתכנים מדעיים שנלמדו בבית הספר. הגישה המובילה במבחני פיזה שמה את הדגש על יכולת התלמיד לחשוב בדרכים מדעיות: לשאול שאלות, לאסוף ראיות, להסביר תופעות, לפרש נתונים ולהעריך טענות במסגרת הקשר שיש בו טענה ונימוקים.

מה חדש במסגרת המושגית בתחום המדעים של פיזה 2025?

3.2

הערכת תחום המדעים של פיזה 2025 מבוססת על הערכות קודמות של הארגון, אך כוללת ארבעה עדכונים שנועדו להתאים את ההערכה למציאות החינוכית והאזרחית בעת הזו.

כשירות החקר ומיומנות פירוש ראיות

בעבר הפריד המחקר בין הערכה של חקר מדעי לבין פירוש נתונים וראיות. בפיזה 2025 שתי היכולות שולבו לכשירות אחת רחבה יותר. השינוי נבע מכך שחקר מדעי ופירוש ראיות הם תהליכים השלובים

זה בזה: איסוף הנתונים נועד לאפשר הסקת מסקנות, ומהן נגזרת קבלת החלטות הנשענות על תהליך של שיפוט עובדות וראיות ופרשנות. כלומר, איסוף הנתונים ופירוש עובדות וראיות הם חלקים מתהליך חקר אחד, לכן תלמיד לא נדרש רק לקרוא גרף או לזהות מסקנה, אלא להבין כיצד הנתונים נאספו, מהן מגבלות אופן המדידה, מה ניתן להסיק מהמצאים וגם מה לא ניתן להסיק מהם.

הדגשת איתור מידע מדעי והשימוש בו

עדכון זה משקף מציאות בה תלמידים נחשפים מדי יום ביומו למידע מדעי ממקורות רבים ומגוונים: אתרי אינטרנט, רשתות חברתיות, פרסומים שונים, ידיעות עיתונאיות, סרטים ופרסומות. כשירות זו דורשת מהתלמיד לא רק לקרוא ולהבין מידע מדעי, אלא להעריך את מהימנות המקור של המידע, לזהות טענות שאינן מבוססות ולהכריע בין פרשנויות מתחרות.

הדגשת הידע הפרוצדורלי והאפיסטמי

בפיזה 2025, לצד הידע ביחס לתוכן המדעי, נמדדת גם הבנת הדרכים שבהן ידע זה נבנה (פרוצדורלי) וההתגבשות של הידע המדעי כמערכת ידע (אפיסטמי). תלמידים נדרשים להבין לא רק מה ידוע במדע, אלא להבין את הידע על המדע וכיצד הוא מתהווה: כיצד שאלות חקר מנוסחות, כיצד מאמתים השערות, כיצד מתמודדים עם אי-ודאות, ומה תפקיד המודלים המדעיים.

הדגשת הקשר בין מדע לקיימות

מבחן פיזה 2025 מדגיש כי אוריינות מדעית כוללת יכולת להבין סוגיות סביבתיות מורכבות, להעריך ידע הנוגע אליהן ולשקול פעולה אישית וקולקטיבית לצורך קבלת החלטות מבוססות ראיות. אין מדובר בהזדהות אידיאולוגית עם קיימות, אלא בבחינת יכולתו של התלמיד להבין מידע מדעי ולפעול באופן מושכל בסוגיות שבהן מדע, חברה וסביבה נפגשים.

המודל הקוגניטיבי - שלוש הכשירויות המדעיות

3.3

המודל הקוגניטיבי של הערכת המדעים במבחן פיזה 2025 מגדיר את מרכיבי האוריינות המדעית ואת הדרך בה הם באים לידי ביטוי במשימות המבחן. כאמור, לאור השינויים שפורטו לעיל, כובד המשקל הוסט למעשה מהערכת יכולות של תכנון ניסויים מדעיים להערכה של יכולות חקר ומיומנויות חשיבה ביקורתית. לכן נקבעו שלוש כשירויות הנדרשות בליבת ההערכה: הסבר תופעות באופן מדעי; תכנון, הערכה ופירוש חקר מדעי; איתור ושימוש במידע מדעי.

הסבר תופעות באופן מדעי

הכשירות הראשונה עוסקת ביכולתם של תלמידים להשתמש בידע מדעי כדי להסביר תופעות טבעיות, טכנולוגיות וסביבתיות. כשירות זו מתבטאת ביכולת לזהות עיקרון מדעי הרלוונטי למצב נתון, להסביר מדוע תופעה מסוימת מתרחשת, להשתמש במודל או בייצוג ויזואלי כדי לתאר תהליך, להעריך הסבר נתון ולזהות אם הוא מתאים לראיות שהוצגו.

משימות בתחום זה עשויות לעסוק בתופעות יומיומיות מוכרות, כגון התחממות חומרים, מעבר חום, נשימה, מחלות, זרימת מים, הפקת אנרגייה ותהליכים אקולוגיים. הן יכולות גם להציג מצבים חדשים שבהם התלמיד נדרש ליישם ידע מדעי בהקשר לא מוכר, כך שהכשירות אינה בוחנת שחזור ידע קודם, אלא את יכולת התלמיד לזהות ידע רלוונטי ולנסח הסבר מבוסס.

הסבר מדעי טוב אינו מסתפק באזכור מושג, אלא קושר בין המושג לבין התופעה. לדוגמה, תלמיד עשוי להידרש להסביר שינוי במצב צבירה באמצעות הקשר בין טמפרטורה, לאנרגייה ולתנועת חלקיקים; לתאר שינוי באוכלוסיית בעלי חיים באמצעות קשרי תלות במערכת אקולוגית; או להסביר תוצאות ניסוי באמצעות הקשר בין משתנים.

תכנון חקר מדעי, הערכה ופירוש ביקורתי של נתונים וראיות

הכשירות השנייה עוסקת בהבנת האופן שבו ידע מדעי מתהווה באמצעות חקר, נתונים וראיות. תלמידים נדרשים להבין מהי שאלת מחקר, כיצד ניתן לבדוק אותה, אילו נתונים נדרשים לשם כך, כיצד יש לאסוף אותם, כיצד לבדוק אם הנתונים אמינים ואיכותיים, וכיצד ניתן לפרש את הנתונים באופן הוגן ומנומק. כשירות זו כוללת שני מרכיבים המשלימים זה את זה. הראשון הוא תכנון חקר: תלמידים נדרשים לזהות את שאלת המחקר, להבחין בין משתנה בלתי-תלוי למשתנה בקרה, להעריך אם ניסוי מתאים לשאלה הנשאלת ולזהות הטיות אפשריות. המרכיב השני הוא פירוש ביקורתי של נתונים: תלמידים נדרשים לזהות מגמה בגרף, להשוות בין קבוצות, להסיק מסקנה מטבלה, להבחין בין מתאם לבין סיבתיות, לזהות חריגים, להעריך את גודל המדגם או להסביר מדוע נדרשת חזרה על מדידה.

איתור, הערכה ושימוש במידע מדעי לקבלת החלטות ופעולה

הכשירות השלישית עוסקת ביכולתם של תלמידים להתמודד עם מידע מדעי כפי שהוא מופיע מחוץ להקשר בית-ספרי פורמלי. תלמידים נדרשים לאתר מידע רלוונטי, להעריך את מהימנות המידע ומקורו, לזהות טענות שאינן מבוססות, ולהשתמש במידע מסוגים שונים כגון גרפים, טבלאות, מפות ונתונים לצורך קבלת החלטות ופעולה. כשירות זו חשובה במיוחד בעידן של מידע מדעי נגיש אך גם של מידע חלקי, מטעה או מוטעה. היא מחברת בין אוריינות מדעית לבין אוריינות מידע, חשיבה ביקורתית ואזרחות דמוקרטית. תלמיד בעל אוריינות מדעית אינו נדרש להפוך למומחה בכל תחום, אך עליו להיות מסוגל להעריך ידע, להבין מהן ראיות מהימנות ולנקוט עמדה מבוססת בנושאים מדעיים הנוגעים לחייו ולחיי קהילתו.

סוגי הידע הנדרשים באוריינות מדעית

3.4

שלוש הכשירויות המדעיות אינן נמדדות בחלל ריק. כדי להסביר תופעה, להעריך חקר, לפרש נתונים ולהשתמש במידע מדעי, תלמיד נדרש להפעיל מספר סוגי ידע. מבחן המדעים של פיזה מבחין בין שלושה סוגי ידע מרכזיים: ידע תוכני, ידע פרוצדורלי וידע אפיסטמי.

ידע תוכני

הידע המדעי על העולם: מושגים, עקרונות, תהליכים ותופעות, רעיונות ותיאוריות בתחומי מדע. בפיזה 2025 ידע זה מאורגן סביב שלוש מערכות עיקריות:

מערכות פיזיקליות – מושגים כגון חומר, אנרגיה, כוחות, תנועה, חשמל, חום ותכונות חומרים.

מערכות חיים – מבנה תאים ואברים, תורשה, אבולוציה, מערכות ביולוגיות ויחסי גומלין בין אורגניזמים לסביבתם.

מערכות כדור הארץ והחלל – אקלים, מזג אוויר, משאבי טבע, מחזורי חומרים, מבנה גאולוגי, כדור הארץ והחלל והשפעת פעילות האדם על מערכות סביבתיות.

חשוב לציין שבפיזה אין די בידע התוכני; התלמיד נדרש ליישם אותו בהקשר מסוים, לקשר אותו לנתונים ולהסיק ממנו מסקנות.

ידע פרוצדורלי

ידע על דרכי החקר המדעי. הוא מתייחס לשאלה כיצד מדענים אוספים ראיות, כיצד מתכננים מחקר, כיצד מודדים, כיצד בודקים השערות וכיצד מעריכים את איכות הנתונים. ידע פרוצדורלי כולל, בין היתר, הבנה של שאלות מחקר טובות, ניסוח השערות, ההבחנה בין משתנה תלוי לבלתי-תלוי, הצורך במשתנים קבועים, חשיבות קבוצות ביקורת, הצורך בחזרות מדידה, תפקיד המהימנות ותקפות המדידה, גודל מדגם מספק וזיהוי הטיות אפשריות.

ידע אפיסטמי

ידע על אופי הידע המדעי: כיצד ידע מדעי נבנה, כיצד הוא נבדק, כיצד הוא מוצדק ומה גבולותיו. ידע אפיסטמי מאפשר לתלמיד להבין שמדע הוא לא אוסף קבוע של עובדות אלא תהליך שיטתי של בניית ידע על בסיס ראיות ונימוקים. בהקשרים אחרים, ידע אפיסטמי מסייע לתלמידים לשאול מי מציג את הטענה ומהו עניינו? אילו ראיות עומדות בבסיסה? האם המסקנה עברה ביקורת עמיתים? האם קיימות פרשנויות חלופיות? והאם אי-הוודאות מוצגת בצורה הוגנת?

רכיב במודל	מה נמדד	דוגמה כללית
הקשר	המצב שבו מוצגת המשימה	בריאות, סביבה, אנרגיה, טכנולוגיה; ברמה אישית, מקומית או גלובלית
ידע	סוג הידע שהתלמיד מפעיל	ידע תוכני, ידע פרוצדורלי, ידע אפיסטמי
כשירות	הפעולה הקוגניטיבית הנדרשת	להסביר תופעות, להעריך חקר ולפרש ראיות, לאתר ולהעריך מידע מדעי
עמדות וזהות	מאפייני רקע ועמדות כלפי מדע	עניין במדע, תפיסת מסוגלות, פעלונות סביבתית

זהות מדעית, עמדות ופעלונות סביבתית

3.5

לצד הכשירויות והידע המדעי, מבחן פיזה 2025 מתייחס גם לרכיבי הזהות המדעית, העמדות כלפי המדע והפעלונות הסביבתית. רכיבים אלה נמדדים בעיקר באמצעות שאלוני רקע שמועברים לצד המבחן, ולא דרך פריטי ההישגים עצמם.

זהות מדעית

זהות מדעית מתייחסת לאופן שבו תלמידים רואים את עצמם ביחס למדע. המחקר מתעניין בשאלות כגון האם התלמיד מתעניין במדע? האם הוא מאמין שהוא מסוגל ללמוד מדעים בהצלחה? האם הוא רואה את המדע כתחום רלוונטי לחייו? והאם הוא יכול לדמיין את עצמו עובד בתחום מדעי בעתיד?

עמדות כלפי מדע

המבחן מתייחס גם לעמדות של תלמידים כלפי מדע ("אמונות אפיסטמיות") ולתפיסות שלהם בנוגע לאופי הידע המדעי. האם תלמידים מבינים שידע מדעי מבוסס על ראיות? האם הם מכירים בכך שמדע הוא תהליך מתפתח שממצאיו עשויים להשתנות? האם הם מסוגלים להבחין בין טענה שנבדקה ונמצאה נתמכת לבין טענה שלא עברה בדיקה מדעית?

פעלונות סביבתית

פעלונות סביבתית הוגדרה במחקר פיזה 2025 כ"תחושת מסוגלות ונכונות לפעול באופן אישי או קולקטיבי לקידום קיימות ולהתמודדות עם סוגיות סביבתיות, על בסיס הבנה מדעית ושיקולים ערכיים". כאמור, הכוונה אינה התחקות אחר עמדות פוליטיות של תלמידים, אלא בחינת המסוגלות שלהם להבין מידע מדעי בנוגע לסוגיות סביבתיות, להעריך אפשרויות פעולה שונות ולשקול את ההשלכות שלהן על בסיס ידע מדעי.

הקשרי המשימות

בחירת ההקשרים שבהם מוצגות המשימות היא סוגיה מרכזית בתכנון מחקר פיזה. מצד אחד, ככל שמגוון ההקשרים עולה, כך כיסוי התחום הנבדק מתרחב ותוקף ההערכה משתפר; מצד שני, הקשרים צריכים להיות נגישים ורלוונטיים לתלמידות ולתלמידים בני 15 ממדינות ותרבויות שונות, כך שלא ייווצר יתרון כלשהו לתלמידים ממדינה מסוימת.

ההקשרים במחקר פיזה 2025 מאורגנים על פי שני צירים: ציר הקרבה לתלמיד - אישי, מקומי-לאומי, גלובלי; וציר תחום הידע המדעי הנדרש.



הקשרים גלובליים

סוגיות הנוגעות לכלל האנושות, לדוגמה שינויי אקלים, מגוון ביולוגי, משאבי טבע, מגפות ואנרגיה.



הקשרים מקומיים ולאומיים

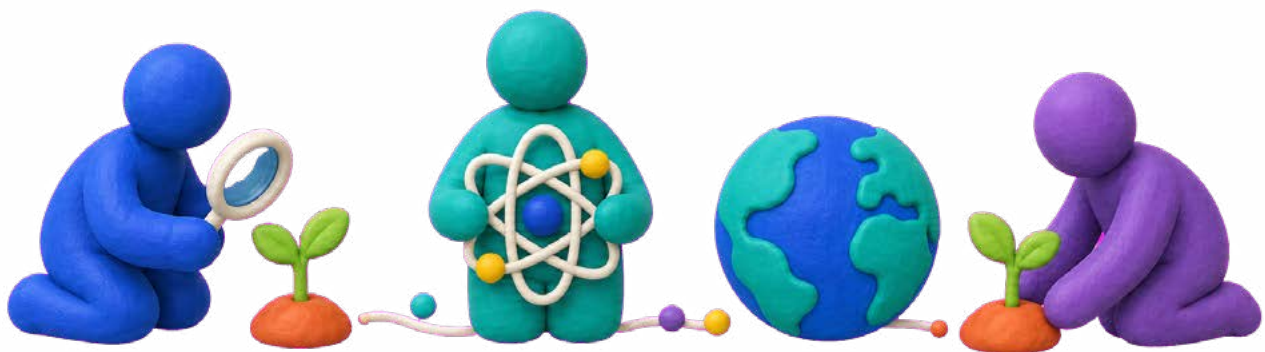
נושאים רחבים כמו איכות אוויר ומים, ניהול פסולת, תחבורה, חקלאות ובריאות הציבור.



הקשרים אישיים

נושאים הנוגעים לתלמיד באופן ישיר כגון בריאות אישית, תזונה, פעילות פנאי, שימוש בטכנולוגיות.

לצד ההקשר, כל משימה כוללת גרָיִין (stimulus) המוצג לתלמיד. הגריין עשוי להיות טקסט, תרשים, טבלה, גרף, תמונה, דיאגרמה, תיאור מצב, אנימציה או סימולציה. לצד הגריין מופיעות שאלה אחת או יותר המתייחסות אליו. לעיתים הגריין כולל את כל המידע הנדרש לתשובה, אך יש שהוא מצריך שילוב בין המידע המוצג לבין ידע מדעי שהתלמיד מביא עימו.



תיאור ההערכה וסוגי הפריטים

4

מבנה יחידות המבחן

4.1

פריטי המדעים מאורגנים ביחידות שכאמור בכל אחת מהן משולב גריין המציג לתלמיד הקשר, מידע, נתונים ותיאור מצב - ולאחר מכן מופיעות שאלות המתייחסות לגריין. יחידה אחת עשויה לכלול שאלות הבוחנות כשירותיות שונות ורמות קושי שונות. מבנה זה מאפשר לתלמיד להעמיק בהקשר מדעי אחד ולגלות יכולת לעבד מידע ברמות שונות של מורכבות.

פרק הזמן הנדרש לכל יחידה נע בין 5 ל-15 דקות, בהתאם למספר הפריטים ולמורכבות הגריין. ביחידות שכללו יותר מפריט אחד מפתחי המבחן הקפידו להפחית את מידת התלות בין הפריטים, כך שאי-הצלחה בפריט אחד לא תשפיע בהכרח על יכולת התשובה בפריט הבא.

סוגי הפריטים

4.2

בתחום המדעים נעשה שימוש בכמה סוגי פריטים המנצלים את יתרונות הסביבה הממוחשבת:

פריטים רבי-ברירה פשוטים - התלמיד בוחר תשובה אחת מתוך מספר אפשרויות. מתאים לבדיקת זיהוי של הסבר, מושג או מסקנה.

פריטים רבי-ברירה מורכבים - התלמיד מסמן את ההיגדים הנכונים לדעתו מתוך רשימה, ולעיתים תיתכן יותר מתשובה אחת נכונה. מתאים לבדיקת הבנה מורכבת יותר.

פריטי גרירה והצבה - התלמיד גורר רכיבים וממקם אותם בטבלה, על גבי תרשים, ציר זמן או מפה מושגית. מתאים לבדיקת הבנה של תהליכים ומבנים.

פריטי סימון בייצוג חזותי - התלמיד מסמן נקודה, אזור או רכיב על גרף, תרשים מערכת, מפה או ציור. מתאים לפרשנות גרפית וויזואלית של נתונים.

פריטי תפריט נפתח - התלמיד בוחר מושג מתאים מרשימה נתונה כדי להשלים משפט מדעי. מאפשר בדיקת ידע תוכן וידע פרוצדורלי ללא הקלדה חופשית.

פריטי סימולציה - התלמיד מפעיל משתנים בניסוי מדומה וצופה בתוצאות. מתאים לבחינת הבנת הקשרים של סיבה ותוצאה ותכנון חקר.

פריטים פתוחים קצרים - התלמיד מקליד הסבר קצר, מסקנה, נימוק או הצעה מנומקת. מאפשר בדיקת הסבר מדעי מלא הכולל קישור בין מושג לתופעה.

כשבעים אחוזים מפריטי המבחן בתחום המדעים קודדו באופן אוטומטי על ידי המערכת, וכשלושים אחוזים היו פתוחים וקודדו בידי בודקים אנושיים. כך התאפשר לתלמידים לנסח הסברים מדעיים מפורטים.

סוג פריט	אופן התגובה	דוגמה כללית
רב-ברירה פשוט	בחירת תשובה אחת	בחירת ההסבר המתאים לתופעה
רב-ברירה מורכב	סימון כמה אפשרויות	סימון אילו מסקנות נתמכות בנתונים
גרירה והצבה	גרירת רכיבים למיקום מתאים	הצבת תוויות על תרשים או סידור שלבים בתהליך
סימון בייצוג חזותי	בחירת אזור, נקודה או רכיב	סימון נקודה בגרף או אזור במערכת ביולוגית
תפריט נפתח	בחירת תשובה מרשימה	השלמת משפט מדעי באמצעות מושג מתאים
סימולציה	הפעלת משתנים וצפייה בתוצאות	בחינת השפעת משתנה בלתי תלוי על תוצאה
מענה פתוח קצר	הקלדת תשובה	ניסוח הסבר מדעי מנומק או מסקנה

עיצוב ופיתוח ההערכה ומחקרי התיקוף

5

פיתוח פריטי המבחן הוא תהליך ממושך ומרובה שלבים. הוא נפרס על פני מספר סבבים שבהם התבקשו המדינות המשתתפות לתת משוב על הפריטים. לאחר מכן, נבדקו הפריטים בידי מומחי לשון ותרגום המייצגים קבוצות שונות של שפות ותרבויות. בדיקות אלה נועדו לאתר בעיות אפשריות בתרגום ולזהות הטיות תרבותיות שעלולות לפגוע באחידות הבחינה ובהשוואה הבינלאומית.

לכן, לצד ההתאמה הלשונית, נציגי המדינות המשתתפות, ובהן ישראל, מתבקשים לסקור כל יחידת מבחן ולהעיר על הקשרים, דוגמאות או מצבים שעשויים להיות זרים, בלתי מוכרים או נתפסים כפוגעניים מבחינה תרבותית לתלמידים במדינתם. לדוגמה, בבתי ספר בחינוך החרדי בישראל הוצאו מהמחקר שאלות שהופנו לתלמידים העוסקות בגילויי אלימות בבית הספר או בניידות חברתית, ומהסקר למנהלים בבתי ספר אלה הוצאו בין היתר שאלות העוסקות בתשתיות דיגיטליות, במוכנות ללמידה דיגיטלית ובשימוש בבית מלאכותית. ממצאי הבדיקה, לצד הנתונים הפסיכומטריים שעולים ממחקר החלוץ, עשויים להוביל להתאמת פריטים או אף לפסילתם במקרים שבהם אותרו אי התאמות תרבותיות, מגדריות או הקשרים שאינם רלוונטיים לחלק מהאוכלוסיות. בישראל, נוסף על הבדיקה הבין-לאומית הכללית, נבחנה התאמת התוכן גם ביחס למגזרים ולזרמי החינוך השונים הפועלים בארץ.

תרגום והתאמה לישראל

5.1

גרסאות המקור הבינלאומיות של המבחן ושאלוני הרקע סופקו באנגלית ובצרפתית. ישראל, בה שתי שפות הערכה רשמיות במחקר (עברית וערבית), ביצעה את תהליכי התרגום וההתאמה לשתיהן. תהליך התרגום כלל:

תרגום כפול עצמאי - כל פריט תורגם בנפרד על ידי שני מתרגמים בלתי תלויים

השוואה ואיחוד - שני התרגומים נותחו, הושו וואוחדו לגרסה אחת

אימות לשוני - הגרסאות הישראליות נבדקו על ידי מאמתיים לשוניים בינלאומיים מטעם ה-OECD לוודא שקילות לשונית לגרסת המקור

עדכון לעומת מחזורים קודמים - פריטים ממחזורי מחקר קודמים שתורגמו בעבר נבחנו מחדש ושנו רק במקרים שבהם אותרו שגיאות או בוצעו עדכונים בגרסת המקור

כל ההתאמות לפריטים, הן בגרסה העברית הן בגרסה הערבית, תועדו בגיליון ההתאמות הייעודי ואושרו על ידי הארגון הבינלאומי.

מחקרי תיקוף על ידי פיזה

5.2

לאחר פיתוח רכיבי המבחן, קיים הארגון הבינלאומי מחקר תיקוף בהשתתפות תלמידים בני 15 ממדינות שונות שנבחנו באב-טיפוס של המבחן. תשובותיהם קודדו בהתאם למחוו הראשוני שפותח. על בסיס הממצאים שעלו מניתוח תשובותיהם, נפסלו פריטים שנמצאו לא מתאימים, ואחרים עודכנו, לרבות תיקון של ניסוחים בלתי ברורים. מחקר תיקוף נוסף נערך לאחר שעודכנו הפריטים והוטמעו השינויים. בשלב זה קודדו תשובות התלמידים בידי בודקים שאינם מומחי תחום, כדי לוודא את היתכנות הבדיקה במחקר החלוץ ולבחון אם ניתן להשיג מהימנות טובה בין שופטים ללא הכשרה מתחום התוכן הנבדק.

מחקר החלוץ בישראל

5.3

מחקר החלוץ של פיזה התקיים כשנה לפני המחקר העיקרי, כחלק מתהליך ההערכות הבינלאומי שנועד לאפשר התאמת המבחן בכל מדינה ובדיקה של התשובות לפי הנוסחים המקומיים, בשפות הרשמיות המקומיות. מטרת מחקר החלוץ כפולה: לספק ראיות לתוקף ולמהימנות של הפריטים בכלל המדינות ובכלל השפות, ולבחון את התהליכים התפעוליים כמו מערך הבדיקה, נוהלי קידוד והעברת הנתונים.

בישראל נערך מחקר החלוץ בחודש מאי 2024 בקבוצת מדגם של בתי ספר המייצגת את האוכלוסיות השונות. ממצאי מחקר החלוץ שימשו לצרכים הבאים:

בחינת תפקוד של כל שאלה, כולל תפקוד קטגוריות התשובות שנקבעו בשלב הפיתוח;

הגדרה מחדש של כל קטגוריות ונושאים שלא פעלו כצפוי;

הוספת דוגמאות נרחבות לתשובות תלמידים כדי להדגים קטגוריות שונות;

פסילת פריטים בודדים שלא הגיעו לאיכות הפסיכומטרית הנדרשת;

ארגון מחדש של יחידות המבחן בהתאם לתחומים, לרמות קושי ולזמן המוקצה לכל יחידה.

הכשרת הבודקים

5.4

לקראת מחקר החלוץ ולקראת המחקר העיקרי נערכו מפגשי הכשרה בינלאומיים לנציגי המדינות המשתתפות. במפגשים אלה הוצגו כללי הקידוד ודוגמאות לתשובות ונדונו מקרי קצה ומצבים עמומים. הערות ותובנות שהועלו במפגשים הוטמעו בשלב שיפור המחוונים.

לאחר ההכשרה הבינלאומית ובהובלת מומחים בכירים בתחומי התוכן, הוכשרו על ידי ראמ"ה בודקים ישראליים לקידוד התשובות בעברית ובערבית. הכשרה זו כללה בדיקת תשובות לדוגמה, מפגשי כיוול בין הבודקים ואיסוף משוב על מצבים שלא קיבלו מענה במחווון הבינלאומי.

הליכי הבדיקה והקידוד

6

העקרונות הכלליים לבדיקת תשובות פתוחות בכל תחומי האוריינות בפיזה דומים. תשובות התלמידים נבדקות בהתאם למחווין בינלאומי מפורט שנועד להבטיח כי תשובות דומות מקודדות באופן דומה בכל המדינות המשתתפות ובכל השפות.

חלק מהפריטים במבחן מקודדים באופן אוטומטי על ידי המערכת, בעיקר פריטים סגורים. פריטים פתוחים נבדקים על ידי בודקים שהוכשרו לכך. הם לא מעריכים את איכות הכתיבה, אלא את המהות המדעית של התשובה ואת מידת הדיוק של ההסבר, כפי שהוגדר במחווין.

שלבי הקידוד

6.1

תהליך קידוד כל שאלה עבר ארבעה שלבים עוקבים:

רלוונטיות: האם תשובת התלמיד מתייחסת לשאלה ולגרין?

מושג: האם התשובה כוללת מושג או עיקרון מדעי רלוונטי?

קישור: האם התשובה קושרת נכון בין המושג לבין התופעה?

שימוש בנתון: כאשר הפריט דורש שימוש בנתונים מהגרין - האם התלמיד אכן השתמש בנתון הנכון לתמיכה בטענתו?

תרשים 1. דוגמה לקידוד פריטים שבדקו חקר, נתונים וראיות



מהימנות בין בודקים

6.2

הקידוד הוא תהליך מורכב המבוסס בחלקו על שיקולי דעת של בודקים, ועל כן יש חשיבות לוודא שקיימת מהימנות גבוהה ביניהם, כלומר מידת הסכמה גבוהה בין בודקים שונים שקידדו את אותה תשובה. הסף המינימלי הנדרש במחקר פיזה הוא 85% הסכמה בין בודקים.

כמו בכל המדינות המשתתפות במחקר כך גם בישראל, כדי לבדוק מהימנות בין בודקים, נדגמו במהלך הבדיקה תשובות שקודדו על ידי כלל הבודקים, ונבדקה מידת ההסכמה ביניהם מבחינת הקצאת ניקוד חלקי, ניקוד מלא או היעדר ניקוד. במקרים שבהם רמת ההסכמה בין הבודקים הייתה נמוכה מ-85% או נמצאו פערים משמעותיים, המומחה הבכיר שב והבהיר את מרחב האפשרויות לפי מחווני הבדיקה והחזיר את השאלות לבדיקה חוזרת לבודקים אצלם נמצא פער מהמחווין.

לצורך בדיקת מהימנות בין מדינות הוקצו לכל פריט תשובות באנגלית, כך שכל מדינה יכולה לקודד אותן לצד התשובות בשפתה. הקידוד של תשובות זהות לכלל המדינות בכל פריט איפשר למפתחי המחקר לוודא שכללי הקידוד מפורשים ומיושמים באופן דומה בכל מדינה ומדינה.

מערכ המחקר בישראל: דגימה, העברה ובקרת איכות

אוכלוסיית היעד בישראל

המחקר העיקרי נערך בישראל בחודשים מרץ - אפריל 2025 במדגם מייצג של אוכלוסיית היעד - תלמידים ותלמידות שגילם נע מ-15 ושלשה חודשים ועד ל-16 וחודשיים הלומדים במוסדות חינוך הנכללים במסגרת הדגימה. על פי כללי הארגון הבין-לאומי, נגרעו ממסגרת זו בתי ספר לחינוך מיוחד בעלי/ות מוגבלויות ובתי ספר ששפת ההוראה בהם אינה עברית או ערבית. כמו כן, בתי הספר במזרח ירושלים אינם נכללים באוכלוסיית המטרה של המחקר הישראלי.

ההשפעה של המצב הביטחוני על ביצוע המחקר

מאז 7 באוקטובר 2023 ישראל מצויה במצב מלחמה. עקב כך, בתיאום עם הארגון הבינלאומי ומנהלי המחקר מטעמו, כבר בהיערכות למחקר החלוץ שהתקיים באביב 2024 הוחרגו מאוכלוסיית המדגם בתי ספר מאזורים שפוננו בדרום הארץ ובצפונה. הגם שביישובי הדרום החלו תושבים לשוב לבתיהם עוד קודם לכן, מתווה החזרה ללימודים באזור הצפון נקבע על ידי משרד החינוך בחודש פברואר 2025, כאשר באופן מעשי שבו מרבית בתי הספר לפעילות באזורים אלה רק לקראת ספטמבר אותה שנה. כיוון שכך, מדיניות ראמ"ה באשר למסגרת המדגם במחקר העיקרי באביב 2025 הייתה שבתי ספר שפוננו בדרום הארץ ובצפונה לא ייכללו ולא ישתתפו במחקר. בתי ספר ששבו לפעילות אך עדיין אחוז גבוה של תלמידים שלומדים בהם לא שבו ללימודים סדירים, לא ישתתפו במבחן. כמו כן, לפי החלטת מחוז צפון הוצאו מהשתתפות במחקר בתי ספר נוספים שמצויים בקו העימות ועקב המצב לא התאפשרה בהם היבחנות. מנהלי בתי הספר שבהם התקיים המבחן יכולים היו לפטור מהשתתפות בו תלמידים אשר הם או קרובי משפחתם מדרגה ראשונה נפגעו במלחמה.

הליך הדגימה הדו-שלבי

הליך הדגימה במחקר פיזה הוא דו-שלבי. בשלב הראשון נדגמים בתי ספר מתוך מסגרת הדגימה הלאומית, תוך שימוש בדגימה מרובדת (stratified sampling) המביאה בחשבון משתנים כמו שפת אוכלוסייה (עברית או ערבית), סוג הפיקוח (ממלכת, ממ"ד או חרדי), רקע חברתי-כלכלי (נמוך, בינוני או גבוה) ועוד. בשלב השני נדגמים תלמידים מתוך בתי הספר שנדגמו והנכללים בהגדרת הגיל במועד הביצוע המתוכנן של המבחן.

גריעה מאוכלוסיית היעד

על כלל התלמידים ובתי הספר שנדגמו להשתתף במחקר. עם זאת, וללא קשר לגריעה מאוכלוסיית המדגם עקב המצב הביטחוני המפורט לעיל על פי תקני פיזה, ניתן לפטור מהמחקר במקרים חריגים מספר מוגבל של תלמידים או מוסדות, ובלבד שבמחקר ישתתפו לפחות 95% מאוכלוסיית היעד.

גריעה ברמת בתי הספר - כללי המחקר מאפשרים לגרוע עד 0.5% מבתי הספר, על ידי הוצאת בתי ספר קטנים מאוד או בתי ספר שאינם נגישים; ועד 2.0% נוספים על ידי הוצאת בתי ספר המיועדים להוראת תלמידים שפטורים מהמבחן, לדוגמה בתי ספר לחינוך מיוחד בלבד או בתי ספר שמלמדים בשפה זרה.

גריעה ברמת תלמידים בתוך בתי ספר - כללי המחקר מאפשרים לגרוע עד 2.5% מהתלמידים שנדגמו במקרים חריגים, דוגמת קיומה של מגבלה קוגניטיבית, התנהגותית או רגשית שאינה מאפשרת השתתפות במבחן; ניסיון מוגבל של התלמיד בשפת המבחן; או כאשר שפת ההוראה בכיתה של התלמיד אינה שפת המבחן.

שכבות הדגימה בישראל

שכבות הדגימה במחקר מוגדרות בתיאום בין ה-OECD לבין ראמ"ה במטרה לשקף את המורכבות והמגוון של מערכת החינוך הישראלית ולאפשר בחינת הקשרים בין משתני רקע לבין רמות הביצוע של התלמידים.

סוג שכבה	שם שכבת הדגימה	רמות בתוך השכבה
גלויה	מגזר שפה	דוברי עברית; דוברי ערבית
	סוג הפיקוח (בקרב דוברי עברית בלבד)	ממלכתי; ממ"ד; חרדי
	מגדר בית הספר (בקרב ממ"ד וחרדי בלבד)	בנות; בנים; מעורב
	מגזר (בקרב דוברי ערבית בלבד)	ערבי; דרוזי; בדואי
	משרד אחראי	משרד החינוך; משרד העבודה, הרווחה והשירותים החברתיים
סמויה	רקע חברתי-כלכלי של בית הספר	גבוה; בינוני; נמוך
	גודל בית הספר	גדול; קטן
	מבנה בית הספר	חט"ב בלבד; תיכון בלבד; מעורב
	מחוז מנהלי גיאוגרפי (לבדואי בלבד)	צפון; דרום

שיטת הדגימה

7.6

בחירת בתי הספר מתוך כל שכבה מתבצעת בשיטת דגימה שיטתית יחסית לגודל (Systematic Probability Proportional to Size), שלפיה הסיכוי של בית ספר להיכלל במדגם עולה ככל שמספר תלמידי פיזה הלומדים בו גדול יותר.

לאחר שנדגם בית ספר, מאותרים בו כלל התלמידות והתלמידים בגיל המתאים להגדרת מחקר פיזה, מתוכם נדגמים בכל בית ספר בין חמישה תלמידים לחמישים ותשעה לכל היותר. שיטת הדגימה הדו-שכבתית גורמת לכך שבית-ספר גדול יותר נדגם בהסתברות גבוהה יותר, אך תלמיד הלומד בו נדגם בהסתברות נמוכה יותר; ולהיפך, בית ספר קטן נדגם בהסתברות נמוכה יותר, אך תלמיד הלומד בו נדגם בהסתברות גבוהה יותר. כך מבטיחים כי ברמה הארצית לכל תלמיד במסגרת הדגימה של פיזה יש סיכוי שווה להיכלל במחקר.

מספר המשתתפים ושיעורי ההשתתפות

7.7

על פי תקני פיזה נדרשים לפחות 150 בתי ספר ו-6,300 תלמידים משתתפים בכל מדינה. שיעור ההשתתפות המינימלי הנדרש הוא **85% ברמת בתי הספר ו-80% ברמת התלמידים**, מתוך אלה שנדגמו ולא נגרעו מהדגימה. בית ספר נחשב משתתף ונתוניו ישוקללו בנתוני המחקר אם לפחות 33% מהתלמידים שנדגמו בו נבחנו בפועל.

כאשר בית ספר שנדגם מסרב להשתתף או מנוע מכך, מאפשרים מארגני המחקר להחליפו בבית ספר חלופי בעל מאפיינים דומים, שנדגם מראש באמצעות אותן שכבות גלויות וסמויות. אשר לקבוצת הבנים החרדים, הוגדל המדגם מראש כדי להבטיח עמידה בשיעורי ההשתתפות הנדרשים בהתאם לכללי המחקר, וזאת לאור ניסיון ממחזורים קודמים לפיו בתי ספר בשכבה זו נוטים שלא לשתף פעולה.

השתתפות בתי ספר בישראל והשתתפות תלמידים

7.8

מתוך 218 בתי ספר שנדגמו בישראל על ידי פיזה מארבעת זרמי החינוך (ממלכתי, ממלכתי-דתי, ממלכתי ערבי, חרדי), השתתפו בפועל 215, שהם 97.7%. 28 בתי ספר מכלל אוכלוסיית המדגם הוחלפו מסיבות שונות בבתי ספר אחרים, שכאמור נדגמו מראש. חשוב לציין כי החל משנת 2009 משתתפות תלמידות במוסדות החינוך החרדי במחקר באופן מלא כאשר שאלות המבחן ושאלוני העמדות עוברים התאמה תרבותית ויעודית (ראו פרק חמש לעיל). במחזור 2025 השתתפו 772 תלמידות מ-23 מוסדות חינוך חרדיים ושיעור השתתפותן בפועל גבוה ועמד על 89%. לעומת זאת, 15 בתי ספר של בנים בחינוך חרדי קיבלו פטור מהשתתפות במחקר (תרבותי – ייחודי) ותלמידים במוסדות אחרים בחינוך החרדי שכן נכללו בדגימה השתתפו בפועל בהיקף נמוך מהנדרש כדי שאפשר יהיה לדווח על ממצאיהם.

סך הכול נדגמו 8,868 תלמידים מתוכם קיבלו פטור פרטני 291, והשתתפו בפועל במבחן 7,330 תלמידים המהווים 85.5%, על פי ההתפלגות המפורטת בטבלה להלן.

שיעור השתתפות מתוך המדגם	השתתפות בפועל		השגות		מסגרת הדגימה*			שכבת דגימה	מגזר שפה
	שיעור	מס' תלמידים	מס' בתי"ס	שיעור מהמדגם	מס' תלמידים	שיעור מכלל האוכלוסייה	מס' תלמידים	מס' בתי"ס	
82%	47%	3,450	94	47%	4,185	46%	66,693	632	ממלכתי
81%	9%	630	19	9%	779	6%	9,069	162	ממ"ד בנות
82%	8%	598	18	8%	726	6%	8,554	182	ממ"ד בנים
64%	2%	125	5	2%	194	2%	2,962	77	ממ"ד מעורב
89%	11%	772	23	10%	872	9%	13,758	525	חרדי בנות
64%	2%	165	11	3%	258	10%	14,183	940	חרדי בנים
70%	0.3%	21	1	0.3%	30	0.1%	93	69	חרדי מעורב
41%	0.2%	14	2	0.4%	34	1%	1,703	37	משרד העבודה, הרווחה והשירותים החברתיים
89%	14%	1,001	27	13%	1,131	12%	18,129	269	ערבי
92%	2%	129	3	2%	140	2%	2,376	27	דרוזי
85%	5%	352	9	5%	415	5%	7,354	92	בדואי
70%	1%	73	3	1%	104	1%	1,524	25	משרד העבודה, הרווחה והשירותים החברתיים
*83%	100%	7,330	215	100%	8,868	100%	146,398	3,037	סך הכול

*נתוני מסגרת הדגימה אינם כוללים בתי ספר שקיבלו פטור מראש, אך הם כוללים תלמידים הזכאים לפטור הלומדים במסגרות. כאמור לעיל, סך התלמידים במדגם (8,868) כולל 291 תלמידים ותלמידות שקיבלו פטור פרטני מסיבות שונות, כך ששיעור ההשתתפות הכולל במחקר העומד על 85.5% מחושב ביחס ל-8,577 התלמידים שבפועל נותרו במדגם. נתוני המדגם בפועל אינם כוללים לא את התלמידים שקיבלו פטור ולא את אלה שלמדו בבתי ספר שלא נבחנו.

שפות ההערכה בישראל

7.9

מבחן פיזה התקיים בישראל בשתי שפות, עברית וערבית. כל תלמיד נבחן בהתאם לשפה שבה מתקיימים הלימודים בבית הספר שבו הוא לומד ולא בהכרח על פי שפת האם שלו, כך שתלמידים במוסדות דוברי עברית נבחנו בעברית, ותלמידים במוסדות דוברי ערבית נבחנו בערבית. תהליכי התרגום, ההתאמה ואימות השקילות הלשונית בין הגרסאות השונות התבצעו בהתאם לנוהלי ה-OECD ובפיקוח גורמי תרגום בינלאומיים מטעם הארגון.

הליך ההיבחנות ואיסוף הנתונים

7.10

כל מנהל/ת בית ספר שנדגם קיבל על כך הודעה מראש כחודשיים לפני מועד המחקר המתוכנן. כל מנהל התבקש למנות איש קשר מטעם בית הספר, בדרך כלל מורה ותיק, רכז שכבה, רכז פדגוגי או סגן מנהל, שתפקידו לתאם את כל ההיבטים התפעוליים של קיום המחקר. המבחנים והשאלונים הועברו על ידי משגיחים חיצוניים בהתאם להנחיות המארגנים הבינלאומיים. בכל בית ספר חולקו התלמידים לשתי כיתות מבחן לפחות, בהתאם לזמינות המחשבים. בכל כיתה מבחן פעלו שני בוחנים – אחד שהתמחה בהפעלת מערכות ממוחשבות ואחד שהתמקד בניהול הליך הבחינה לפי הנחיות בינלאומיות – לצד נוכחות מורה מטעם בית הספר ששמר על הסדר, ללא כל מעורבות בהליך עצמו. משך הבחינה היה כשעתיים. כל גרסת מבחן נפתחה בפרק הנחיות כללי להיכרות עם ממשק המבחן הממוחשב, ולאחריו החל המבחן עצמו המורכב משני פרקים משני תחומי תוכן שונים. בסיום המבחן יצאו התלמידים להפסקה של כ-20 דקות, ולאחריה שבו וענו על שאלון שנמשך כ-35 דקות.

ניהול הנתונים, אבטחה ושמירה

7.11

הנתונים מכל בתי הספר שנבחנו הועברו באופן מקוון למרכז המחקר הבינלאומי. כלל נתוני הבדיקה – הן אלה שקודדו אוטומטית, הן אלה שקודדו בידי בודקים אנושיים – הועברו ישירות למרכז הבינלאומי לחישוב הציונים ולעיבוד המלא של הנתונים. ראש"ה פעלה לאורך כל שלבי הדגימה, ההיבחנות והעברת הנתונים בהתאם למדיניות אבטחת המידע של ACER (המפעיל הבינלאומי) ולנוהלי ה-OECD להגנת פרטיות. כל הנתונים הוצפנו בהעברה ובאחסון, ורשימות הקישור בין שמות התלמידים למזהי המחקר נשמרו אצל ראש"ה בלבד ולא הועברו לפלטפורמה הבינלאומית.

בקרת איכות על ביצוע המבחן

7.12

ראש"ה פעלה במישורים שונים כדי להבטיח את ביצועו התקין של המבחן בבתי הספר שנדגמו, לפי כל אמות המידה המקצועיות שנקבעו על ידי הארגון הבינלאומי. בקורות אלה כללו:

סודיות חומרי המבחן מניעת חשיפה של פריטים לפני המבחן ואחריו	עמידה בתסריט ההעברה וידוא שהמשגיחים פועלים בהתאם לנוהל הסטנדרטי שהוגדר	הכשרת המשגיחים וידוא שכל משגיח עבר הכשרה מלאה ועמד בדרישות המוגדרות
נוכחות התלמידים תיעוד שיעורי ההשתתפות ורישום תלמידים שנעדרו	תקלות מענים במהלך ההיבחנות על פי הנחיות הארגון	תקינות המערכת בדיקות שגרתיות של פלטפורמת המבחן לפני כל יום בדיקה

במקביל, הארגון הבינלאומי מינה מפקחים מטעמו שתפקידם היה לערוך ביקורות בבתי ספר שנדגמו ולבדוק את מוכנותם לביצוע המבחן על פי הנהלים. המפקחים בדקו גם את אופן ביצוע המבחן בפועל ביום הבחינה, לרבות תפקוד המשגיחים, ההיבטים התפעוליים וצוות בית הספר. הם דיווחו ישירות לצוות פיזה הבינלאומי וככל שנמצאו חריגות הן דווחו לראש"ה.

שאלוני הרקע, מערך המדידה והניתוח הסטטיסטי

8

מדד חברתי-תרבותי-כלכלי של פיזה

8.1

ESCS (Index of Economic, Social and Cultural Status) - הוא המדד הבינלאומי המחושב על סמך דיווחים העצמיים של התלמידים בשאלוני הרקע והקשרי המשימות. המדד מורכב משלושה רכיבים מרכזיים המשוקללים באמצעות ניתוח רכיבים עיקריים (PCA- Principle Component Analysis):

מעמד תעסוקתי של ההורים: הסטטוס התעסוקתי הגבוה ביותר של אחד משני ההורים, המקודד על פי סולם הסטטוס התעסוקתי הבינלאומי האחד למקצועות (ISCO).

השכלת ההורים: שנות הלימוד והדרגה החינוכית/אקדמית הגבוהה ביותר של אחד ההורים, המותאמת לסולם הדירוג הבינלאומי האחד להשכלה ISCED.

משאבי הבית: מדד מצרפי הבוחן את זמינותם של עושר חומרי, נכסים טכנולוגיים ומשאבים תרבותיים בביתו של התלמיד, ובכלל זה קיומם של שולחן כתיבה אישי, מקום שקט ללימודים, מחשב, תוכנות לימודיות, ספרי עזר, ומספר הספרים הכולל בבית (המשמש כמשתנה מתווך חזק לסביבה אוריינית).

מדד ה-ESCS נקבע כסולם סטנדרטי רציף בעל ממוצע 0 וסטיית תקן של 1 בקרב כלל מדינות ה-OECD. מדד מצרפי זה הוא כלי מרכזי להערכת מידת השוויוניות במערכת החינוך הלאומית ולבחירת עוצמת הקשר שבין הרקע המשפחתי לבין הישגיו הלימודיים של התלמיד.

מערך המדידה, הכיול והניתוח הפסיכומטרי

8.2

כדי לתרגם את תגובותיהם הגולמיות של התלמידים במבחן הקוגניטיבי ובשאלוני הרקע לסולמות מדידה בני-השוואה ונטולי הטיות, נשען מחקר פיזה על מערך מורכב של מודלים פסיכומטריים ושיטות אמידה מתקדמות. להלן פירוט העקרונות המתודולוגיים שיושמו בעיבוד הנתונים של מחזור 2025.

מערך משקלות הדגימה ותיקוני אי-השבה

מחקר פיזה מבוסס על תכנון דגימה מרובדת דו-שלבית: בשלב הראשון נדגמים בתי ספר מתוך שכבות אפיון מוגדרות (Explicit Strata), ובשלב השני נדגמים באקראי תלמידים בני 15 מתוך רשימות התלמידים המעודכנות במוסדות החינוך שנבחרו. בשל שיטה זו, סיכויי ההיכללות של תלמיד מסוים במדגם אינם אחידים והם משתנים בראש ובראשונה בהתאם לגודל בית הספר שבו הוא לומד.

כדי לתקן את העיוות הזה, לכל תלמיד שנבחר מוצמד "משקל דגימה" מיוחד, שהוא למעשה ההיפוך המתמטי של סיכויי ההגרלה שלו. כך תלמיד שהיה לו סיכוי קטן מאוד להידגם (למשל, מכיוון שהוא לומד בבית ספר קטן ומרוחק) יקבל משקל גדול יותר במערכת הנתונים, מכיוון שהתשובות שלו צריכות לייצג קבוצה רחבה של תלמידים דומים לו באוכלוסייה שלא נבחנו בפועל. לעומת זאת, תלמיד שהיה לו סיכוי גבוה להידגם יקבל משקל קטן יותר כדי לשמור על איזון מדויק והוגן ולא לייצר תמונה מעוותת של כלל תלמידי ישראל בני ה-15. משקולת בסיס זו משולבת עם "מכפיל תיקון לאי-השבה" (Non-response Adjustment Factor) שנועד לפצות על אובדן נתונים ברמת בית הספר או ברמת התלמיד ולמנוע הטיות ייצוג.

בישראל מוגדרת שכבת דגימה נפרדת עבור בתי הספר החרדיים לבנים, אוכלוסייה המאופיינת היסטורית בשיעורי השתתפות נמוכים. משקלי הדגימה בשכבה זו חושבו בהליך התקני של PISA: משקל בסיס ובעקבותיו התאמות לאי-היענות ברמת בית הספר וברמת התלמיד. בשל ההשתתפות והכיסוי הנמוכים יחסית, המשיבים בשכבה נושאים משקל סופי גבוה מן הממוצע הארצי. השקלול משיב לשכבה את חלקה היחסי באוכלוסיית התלמידים, כך שהאומדנים הארציים נותרים בלתי-מוטים בקירוב, בכפוף להנחה שאי-ההיענות מתפלגת באקראי בהינתן משתני השקלול. עם זאת, המשקלים הגבוהים מגדילים את תרומת השכבה לשונות הדגימה ומעצימים את השפעתם של מעט משיבים על האומדן, וכל הטיה שיורית הנובעת מאי-היענות או מכיסוי חסר בשכבה זו מצוינת כמגבלה.

קישור בין מחזורים

אחת המטרות המרכזיות של מחקר פיזה היא אומדן של מגמות שינוי בהישגי מערכות החינוך לאורך זמן. השוואת הציונים בין המחזורים השונים מתאפשרת באמצעות שילובם של פריטי עוגן – אלה משימות מבחן קבועות המועברות באופן זהה במחזורי מחקר עוקבים. תגובות התלמידים מנותחות באמצעות מודלים של תיאוריית התגובה לפריט (IRT – Item Response Theory), המאפשרים לכייל את ציוני המחזור הנוכחי ישירות אל סולם המקור ההיסטורי של כל תחום דעת.

עם זאת, תהליך הקישור הסטטיסטי כרוך באי-ודאות מתמטית מובנית, המכונה "שגיאת קישור" (Linking Error). שגיאה זו נובעת מכך שהקישור בין הסולמות נשען על מדגם מוגדר ומצומצם של פריטי עוגן משותפים ולא על מרחב הפריטים הפוטנציאלי המלא. מחזור פיזה 2025 מאופיין ברגישות פסיכומטרית גבוהה במיוחד בהיבט זה, עקב עדכון סולם האוריינות המדעית ב-2025 והגדרתו מחדש כתחום המרכזי של המחקר (בדומה לעדכון המבני המקיף שנערך לסולם המדעים במחזור 2015). לפיכך, בעת בחינת פערים ומגמות רב-שנתיות, שגיאת הקישור מובאת בחשבון באופן פורמלי במסגרת חישוב טעויות התקן של השינויים, וזאת כדי למנוע קביעה שגויה של מובהקות סטטיסטית להבדלים קלים הנובעים מתנודות כיוול.

בניית מדדי השאלון בשיטת WLE

רבים ממשתני הרקע, העמדות וההקשרים של התלמידים (כגון מסוגלות עצמית, תפיסת אקלים כיתתי או הנאה מלימודי מדעים) הם מושגים מופשטים וסמויים שלא ניתן למדוד באמצעות שאלה בודדת. לצורך בניית מדדים רציפים ומורכבים משאלוני הדיווח העצמי, נעשה שימוש בשיטת אומדן נראות מקסימלית משוקללת.

שיטה זו מבוססת על מודל רש (Rasch Model) ומייצרת עבור כל תלמיד ציון מיקום אופטימלי על פני הרצף הלסנטי של הממד. יתרונה המרכזי של השיטה טמון ביכולתה לתקן את הטיות האומדן המאפיינות אומדני נראות מקסימלית רגילים כאשר מדובר במספר קטן יחסית של היגדים או שאלות. המדדים המופקים מיוצבים ומנורמלים כך שממוצע מדינות ה-OECD נקבע על 0 וסטיית התקן על 1, דבר המאפשר לחוקרים לבצע ניתוחי מתאם ורגרסיה ישירים בין עמדות התלמידים להישגיהם הקוגניטיביים.

אומדן היכולת והפקת ערכים סבירים

בשל היקפו הגדול של מרחב הפריטים במחקר פיזה, לא ניתן להעביר את כלל משימות המבחן לכל תלמיד בלי לייצר עומס בחינה בלתי סביר. לפיכך, המחקר משתמש בעיצוב דגימת מטריצה, שבו כל תלמיד משיב רק על חלק קטן ומסוים מתוך תמהיל הפריטים והנוסחים הכולל. במצב זה, הפקה של ציון נקודתי יחיד עבור תלמיד בודד מלווה בטעות מדידה גדולה מדי אשר מובילה לתת-אמידה משמעותית של השונות באוכלוסייה ולהטיות של ממש בחישוב מדדי פיזור (כגון סטיות תקן, אחוזונים ושיעורי תלמידים ברמות ברורות).

כדי להתגבר על מגבלה מתודולוגית זו, החוקרים מיישמים את טכניקת הערכים הסבירים (Plausible Values – PV) המבוססת על גישה בייסיאנית רב-ממדית. במקום לחשב ציון קבוע, מודל ה-IRT משלב את נתוני התגובות לקובצי המבחן יחד עם מערך עצום של משתני רקע דמוגרפיים מתוך השאלונים כדי לאמוד את פונקציית ההתפלגות המותנית של היכולת עבור כל תלמיד. מתוך התפלגות מותנית זו, התוכנה דוגמת באקראי עשרה ערכים סבירים (PVs 10) לכל נבחן (במחזורים שעברו נעשה שימוש בחמישה ערכים בלבד). חישוב של כל אומדן סטטיסטי רשמי (כגון ממוצע לאומי או פער מגזרי) מבוצע באופן עצמאי לחלוטין, עשר פעמים בנפרד – פעם אחת עבור כל אחד מעשרת ה-PVs. האומדן הסופי המדווח בדוח מחושב כממוצע האריתמטי הפשוט של עשרת החישובים הללו. בכך מובטחת אמידה מדויקת יותר, יציבה ונטולת הטיות של התפלגות הישגי האוכלוסייה במערכת החינוך.

אומדן אי-ודאות וטעויות תקן בשיטת BRR

מערך הדגימה המורכב של פיזה המבוסס על דגימת אשכולות מרובדת (בתי ספר כשלב ראשון), מפר את הנחת היסוד של דגימה מקרית פשוטה (SRS – Simple Random Sampling). שימוש בנוסחאות סטטיסטיות קלאסיות לחישוב טעויות תקן יוביל לתת-אומדן משמעותי של אי-הוודאות (בשל הדמיון

הפנימי הקיים בין תלמידים הלומדים באותו מוסד חינוכי) ויגרום לקביעה שגויה של מובהקות סטטיסטית לפערים מקריים.

לשם אומדן מהימן של טעויות התקן, המחקר מיישם שיטת שכפול רפלקטיבית מתקדמת, המכונה "שכפול רפליקטיבי מאוזן חוזר" (Balanced Repeated Replication - BRR) עם התאמת פיי (Fay's method). במסגרת שיטה זו, מסד הנתונים הלאומי של בתי הספר מחולק ל-80 שכבות אפיון מדומות, ומתוכן מופקות 80 משקולות שכפול (Replicate Weights) שונות עבור כל רשומת תלמיד, כאשר כל משקל משקף וריאציה מבוקרת של המדגם המקורי.

כל נתון סטטיסטי או פרמטר לאומי מחושב 80 פעמים עבור כל אחת ממשקלות השכפול, והשונות בין 80 האומדנים הללו משקפת באופן מדויק את אי-הוודאות הנובעת ממערך הדגימה בשטח. טעות התקן הסופית המדווחת בדוח משלבת באופן פורמלי הן את שונות הדגימה (המופקת מ-80 השכפולים) והן את שונות המדידה (הנובעת מהשימוש בעשרת הערכים הסבירים), בהתאם לנוסחה הבינלאומית המקובלת של מחקרי ה-OECD.

רמות הבקיאות

9

סולם ציוני האוריינות בתחומים השונים פותח על ידי הארגון הבינלאומי, ובו נקבע בשנת הבסיס כממוצע מדינות ה-OECD הציון 500 נקודות עם סטיית תקן של 100. כדי להקנות לציונים פירוש מהותי שמעבר למשמעות המספרית, חולק הסולם לרמות בקיאות שכל אחת מהן מתארת בצורה מוגדרת ומפורטת את יכולות האוריינות המדעית האופייניות לתלמידים הממוקמים בה.

במחזור הנוכחי מוגדרות שבע רמות בקיאות בתחום האוריינות המדעית: רמות 2 עד 6 כאשר רמה 2 הוגדרה כרמת הבסיס ואילו רמה 5 ורמה 6 מוגדרות כמצטיינים. תת-חלוקה זו של רמה 1 מאפשרת אבחנה מעמיקה ומפורטת יותר בין תלמידים מתקשים. בסולם רמות הבקיאות באוריינות קריאה ובאוריינות במתמטיקה מחולקת רמת המתקשים לשלוש תת-רמות: 1א, 1ב, 1ג. (בהתאמה בטבלאות להלן). תלמידים שהשיגיהם מתחת לרמות אלה מוגדרים על ידי פיזה "מתחת לסף".

רמות הבקיאות באוריינות מדעית



9.1

רמה	תיאור היכולות האופייניות לרמה
6	תלמידים ברמה זו מסוגלים לזהות, להסביר וליישם ידע מדעי מורכב ורעיונות בהקשרים מגוונים ומורכבים. הם מסוגלים לקשור מקורות מידע ממגוון תחומים, להעריך בצורה ביקורתית פרשנויות מתחרות ולגבש עמדה מנומקת ומבוססת ראיות. הם מסוגלים להעריך עיצובי חקר מדעי, לזהות מגבלות מתודולוגיות ולהציע שיפורים. הם מפגינים חשיבה יצירתית ורפלקטיבית גם בהקשרים מדעיים שאינם מוכרים.
5	תלמידים ברמה זו מסוגלים לזהות את הרכיבים המדעיים של מצבים מורכבים וליישם מושגים ומידע מדעי לשאלות שאינן מוכרות. הם מסוגלים להשוות נתונים ממקורות שונים ולגבש הסברים מדעיים מנומקים. הם מזהים הנחות, ראיות ונימוקים בתוך טקסטים מדעיים, ומסוגלים לפתח טיעון קוהרנטי המקשר בין ראיות לטענות.
4	תלמידים ברמה זו מסוגלים לפעול באופן אפקטיבי במצבים הכוללים תופעות שעשויות להיות מוכרות אך דורשות הסקה. הם בוחרים ומשלבים הסברים ממקורות שונים ומדיסציפלינות מדעיות שונות וקושרים הסברים אלה ישירות להיבטים ספציפיים של המצב. הם מפרשים נתונים מגרפים ומטבלאות ומנמקים הסברים מדעיים על בסיס קשרים סיבתיים.
3	תלמידים ברמה זו מסוגלים לזהות בעיות מדעיות שהוגדרו בצורה ברורה ולהסביר תופעות מוכרות על בסיס ידע מדעי בסיסי. הם בוחרים בין הסברים אפשריים ותומכים בטענות בעזרת ראיות פשוטות. הם מפרשים נתונים ישירים מגרפים ומסכמים ממצאים.
2 רמת הבסיס	תלמידים ברמה זו מסוגלים להשתמש בידע מדעי בסיסי ומוכר לצורך מתן הסברים תקינים, אם כי לא עמוקים. הם מזהים מאפיינים של חקר מדעי פשוט ומסיקים מסקנות ישירות מנתונים פשוטים. זוהי רמת הבסיס של האוריינות המדעית.
1א	תלמידים ברמה זו מסוגלים לזהות הסברים מדעיים פשוטים בהקשרים מוכרים מאוד ולהבחין בין תופעות מדעיות לתופעות לא מדעיות. הם מסיקים מסקנות ישירות מאוד מנתונים ספציפיים.
1ב	תלמידים ברמה זו מסוגלים להפעיל ידע מדעי מאוד בסיסי ולהסיק מסקנות ישירות מראיות פשוטות מאוד.



רמות הבקיאות באוריינות קריאה

רמה	תיאור היכולות האופייניות לרמה
6	קוראים ברמה זו מסוגלים להבין טקסטים ארוכים ומופשטים שבהם המידע הרלוונטי טמון עמוק בתוך הטקסט ואינו קשור ישירות למשימה. הם משווים, מנגידים ומשלבים מידע המייצג נקודות מבט מרובות ולעיתים סותרות, על בסיס קריטריונים מרובים, ומייצרים הסקות בין קטעי מידע מרוחקים. הם מסוגלים להעריך בצורה ביקורתית את מקור הטקסט ביחס לתוכנו, להשוות בין טקסטים ולזהות ולפתור סתירות בין-טקסטואליות באמצעות הסקות לגבי מהימנות המקורות ואינטרסים גלויים וסמויים.
5	קוראים ברמה זו מסוגלים להבין טקסטים ארוכים ולזהות מידע רלוונטי גם כאשר הוא עלול להיעלם מן העין. הם מבצעים הסקות סיבתיות ואחרות על בסיס הבנה עמוקה של קטעי טקסט מורחבים, ועונים על שאלות עקיפות בהסיקם את הקשר בין השאלה לבין מידע הפזור ברחבי מספר טקסטים. הם מבחינים בין תוכן לבין מטרה, בין עובדה לבין דעה ומסוגלים להעריך ניטרליות והטיה בהסתמך על רמזים מפורשים ומרומזים.
4	קוראים ברמה זו מסוגלים להבין קטעי טקסט מורחבים, בודדים ומרובים. הם מפרשים משמעות של ניואנסים לשוניים בחלק מהטקסט תוך התחשבות בטקסט כולו. הם משווים נקודות מבט ומסיקים מסקנות ממקורות מרובים. הם מאתרים ומשלבים מספר קטעי מידע בנוכחות מסיחים מתעתעים, ומסוגלים להעריך את הקשר בין היגדים ספציפיים לעמדת הכותב הכוללת.
3	קוראים ברמה זו מייצגים את המשמעות המילולית של טקסטים בודדים או מרובים ללא רמזים מפורשים. הם משלבים תוכן ומייצרים הסקות בסיסיות ומתקדמות, ומסוגלים לזהות את הרעיון הראשי, להבין קשר או לפרש מילה וביטוי כאשר המידע הנדרש מצוי בעמוד אחד. הם מאתרים מידע על בסיס הוראות עקיפות ומשווים נקודות מבט של כותבים שונים.
2 רמת הבסיס	קוראים ברמה זו מזהים את הרעיון הראשי בטקסט באורך מתון. הם מבינים קשרים ומפיקים משמעות בתוך חלק מוגבל של הטקסט, גם כאשר המידע אינו בולט לעין, ומאתרים פיסות מידע אחת או יותר על פי קריטריונים מרובים ולעיתים מרומזים. הם מסוגלים להשוות טענות ולהעריך את הנימוקים התומכים בהן על בסיס היגדים קצרים ומפורשים. זוהי רמת הבסיס של אוריינות הקריאה.
1א	קוראים ברמה זו מבינים את המשמעות המילולית של משפטים או קטעים קצרים. הם מסוגלים לזהות את הנושא המרכזי או את מטרת הכותב בטקסט על נושא מוכר, ולאתר פיסות מידע עצמאיות, אחת או יותר, בתוך טקסטים קצרים. רוב המשימות ברמה זו כוללות הוראות מפורשות לגבי מה יש לעשות, כיצד לעשות זאת, ולכן יש לפנות בטקסט.
1ב	קוראים ברמה זו מסוגלים להעריך את המשמעות המילולית של משפטים פשוטים ולפרש את המשמעות המילולית של טקסטים באמצעות קשרים פשוטים בין פיסות מידע סמוכות. הם מאתרים פיסות מידע אחת, בולטת ומפורשת, במשפט יחיד, בטקסט קצר או ברשימה פשוטה. הטקסטים ברמה זו קצרים ומספקים תמיכה לקורא, כגון חזרות, תמונות או סמלים מוכרים.
1ג	תלמידים ברמה זו יכולים להבין את משמעותם המילולית של משפטים קצרים ופשוטים מבחינה תחבירית, ולקרוא לצורך מטרה פשוטה וברורה במשך זמן מוגבל



רמות הבקאות באוריינות מתמטית

רמה	תיאור היכולות האופייניות לרמה
6	תלמידים ברמה זו מסוגלים לפתור בעיות מופשטות ומפגינים חשיבה יצירתית וגמישה. הם מזהים מתי ניתן ליישם פרוצדורה שאינה מפורטת בבעיה בהקשר לא סטנדרטי, ומסוגלים לקשר מקורות מידע וייצוגים שונים - כולל שימוש יעיל בסימולציות וגיליונות אלקטרוניים. הם מפגינים יכולת חשיבה ביקורתית ושליטה בפעולות מתמטיות פורמליות, ומסוגלים לתקשר את נימוקיהם בצורה ברורה ולהעריך את התאמת פתרונם להקשר המקורי.
5	תלמידים ברמה זו מפתחים מודלים למצבים מורכבים ועובדים עימם, מזהים אילוצים ומגדירים הנחות. הם מיישמים אסטרטגיות פתרון בעיות שיטתיות ומתוכננות היטב למשימות מאתגרות יותר - כגון תכנון ניסוי, עיצוב פרוצדורה אופטימלית או עבודה עם ויזואליזציות מורכבות. הם פותרים בעיות שפתרוןן כולל לעיתים ידע מתמטי שאינו מצוין מפורשות במשימה, ומסוגלים לבחון את תוצאותיהם ביחס להקשר המציאותי.
4	תלמידים ברמה זו הנדרשים להתמודדות עם מצבים מוחשיים מורכבים, עובדים בצורה אפקטיבית עם מודלים מפורשים הכוללים לעיתים שני משתנים, ומסוגלים לעבוד גם עם מודלים לא מוגדרים באמצעות חשיבה חישובית. הם מתחילים לעסוק בהיבטי חשיבה ביקורתית כגון הערכת סבירות תוצאה, ומסוגלים לבחור ולשלב ייצוגים שונים - כולל ייצוגים סמליים וגרפיים. הם בונים ומתקשרים הסברים וטיעונים על בסיס פירושיהם, נימוקיהם ומתודולוגיית הפתרון שלהם.
3	תלמידים ברמה זו מגבשים אסטרטגיות פתרון, כולל אסטרטגיות הדורשות קבלת החלטות רצופה או גמישות בהבנת מושגים מוכרים. הם מתחילים להשתמש במיומנויות חשיבה חישובית לפיתוח אסטרטגיית פתרון. הם פותרים משימות הדורשות ביצוע מספר חישובים שגרתיים שונים שאינם מוגדרים כולם בבירור בניסוח הבעיה. הם מסוגלים לפרש ייצוגים ממקורות מידע שונים ולנמק מהם ישירות.
2 רמת הבסיס	תלמידים ברמה זו מזהים מצבים שבהם עליהם לתכנן אסטרטגיות פשוטות לפתרון בעיות, כולל הרצת סימולציות פשוטות הכוללות משתנה אחד. הם מוציאים מידע רלוונטי ממקורות שונים המשתמשים באופני ייצוג מורכבים יותר, כגון טבלאות דו-ממדיות, תרשימים וייצוגים דו-ממדיים של אובייקטים תלת-ממדיים. הם מפגינים הבנה בסיסית של יחסים פונקציונליים ופותרים בעיות הכוללות יחסים פשוטים. זוהי רמת הבסיס של האוריינות המתמטית.
1א	תלמידים ברמה זו עונים על שאלות הכוללות הקשרים פשוטים שבהם כל המידע הנדרש נמצא לפניהם והשאלות מוגדרות בבירור. הם מסוגלים לבצע פרוצדורות פשוטות ושגרתיות על פי הוראות ישירות במצבים מפורשים. הם עוסקים בפעולות שהן ברורות או הדורשות סינתזה מינימלית ביותר, ומפעילים אלגוריתמים, נוסחאות, פרוצדורות או מוסכמות בסיסיים לפתרון בעיות הכוללות בדרך כלל מספרים שלמים.
1ב	תלמידים ברמה זו עונים על שאלות הכוללות הקשרים קלים להבנה שבהם כל המידע הנדרש ניתן בצורה ברורה בייצוג פשוט (טבלאי או גרפי), ומזהים מתי מידע מסוים מיותר לצורך השאלה הספציפית. הם מבצעים חישובים פשוטים עם מספרים שלמים הנובעים מהוראות ברורות המוגדרות בטקסט קצר ופשוט תחבירית.
1ג	ברמה זו התלמידים יכולים לענות על שאלות שההקשר שלהן קל להבנה, וכל המידע הדרוש למענה נתון ומוצג באמצעות ייצוגים פשוטים (כגון, טבלה קטנה או תמונה) לפי הנחיות מפורשות המוגדרות בצורה תחבירית קצרה ופשוטה. הם מסוגלים לעקוב אחרי הנחיה שמתארת שלב יחיד או פעולה יחידה.

10 נספח | מונחים מרכזיים

מונח	הגדרה
אוריינות מדעית	היכולת להשתמש בידע מדעי, בדרכי חקר ובמידע מבוסס ראיות כדי להסביר תופעות, להעריך טענות ולקבל החלטות מושכלות.
גִּרְיִין (stimulus)	חומר פתיחה המוצג לתלמיד, כגון טקסט, תרשים, טבלה, גרף, נתונים, דיאגרמה או סימולציה, שעל בסיסו מופיעות שאלה אחת או יותר.
ידע תוכן	ידע על מושגים, עקרונות ותהליכים במדעים (מערכות פיזיקליות, חיים, כדור הארץ והחלל).
ידע פרוצדורלי	ידע על דרכי חקר מדעי: תכנון ניסויים, מדידה, משתנים, קבוצות ביקורת ואיסוף ראיות.
ידע אפיסטמי	ידע על אופי הידע המדעי: כיצד הוא נבנה, כיצד מוצדק, מה תפקיד הראיות ואי-הוודאות.
כשירות מדעית	אחת משלוש הכשירויות הנמדדות בפיזה: הסבר תופעות; בנייה והערכה של חקר ופירוש ראיות; ואיתור והערכה של מידע מדעי.
פעלנות סביבתית	תחושת מסוגלות ונכונות לפעול באופן אישי או קולקטיבי לקידום קיימות ולהתמודדות עם סוגיות סביבתיות על בסיס מדעי.
גרירה והצבה (drag-and-drop)	פורמט תגובה ממוחשב שבו התלמיד גורר רכיב וממקם אותו במקום מתאים.
מחקר חלוץ (Field Trial)	שלב ראשון של איסוף נתונים שנועד לבחון את איכות הפריטים ואת תקינות ההליכים התפעוליים, כשנה לפני המחקר העיקרי.
מחקר עיקרי (Main Survey)	המחקר המרכזי שנועד לייצר נתונים השוואתיים בינלאומיים לאחר עיבוד ממצאי מחקר החלוץ.
מהימנות בין-בודקים	מידת ההסכמה בין בודקים שונים שקידדו את אותן תשובות; מבטאת עקביות ואמינות של תהליך הקידוד. הסף הנדרש בפיזה: 85%.
דגימה מרובדת (Stratified Sampling)	שיטת דגימה שבה האוכלוסייה מחולקת לשכבות (מגזר, אזור, גודל בית ספר), ומכל שכבה נדגמת כמות מוגדרת.
LDW (Learning in the Digital World)	תחום חדשני בפיזה 2025 העוסק במיומנויות ובעמדות הנדרשות ללמידה עצמאית ולאורך חיים בסביבות דיגיטליות.
PISA Quality Monitor (PQM)	מפקח פיזה בינלאומי המבקר בבתי ספר במהלך ההיבחנות כדי לוודא שתנאי ההיבחנות עומדים בסטנדרטים.
ראמ"ה	הרשות הארצית למדידה והערכה בחינוך האחראית על ביצוע מחקר פיזה בישראל, על תרגום המבחן, הדגימה, הכשרת הבודקים ועיבוד הנתונים.
OECD	הארגון לשיתוף פעולה ולפיתוח כלכלי. הגוף הבינלאומי האחראי על ניהול מחקר פיזה ופרסום ממצאיו.
ACER	Australian Council for Educational Research – הקבלן הבינלאומי הראשי של פיזה 2025 האחראי על עיצוב, פיתוח ויישום המחקר.

ספמבר 2026, אלול תשפ"ו



משרד החינוך