



פירלס 2016

PIRLS – Progress in International Reading Literacy Study

מחקר בין-לאומי להערכת אוריינות קריאה
של תלמידי כיתה ד'



מבט ישראלי

פרק 3 – שיטת המחקר

הדוח המלא של מחקר פירלס 2016 מפורסם באתר הראמ"ה

פרק 3: שיטת המחקר

ביצוע מחקר פירלס, החל בשלב ההיערכות לקראתו דרך איסוף המידע וניתוח התוצאות ועד פרסומן, אורך כחמש שנים. ההיערכות לקראת המחקר החלה בשנת 2013 במרכז המחקר הבין-לאומי. בשלב ראשון עודכנו המסגרת המושגית ופותחו כלי ההערכה (מבחן פירלס להערכת קריאה של טקסטים מודפסים, מבחן אי-פירלס להערכת קריאה בסביבה ממוחשבת, ואף השאלונים לתלמידים, למורים לשפה, למנהלי בתי-הספר ולהורי התלמידים). עם סיומו של תהליך הפיתוח העביר מרכז המחקר את החומרים שפותחו למדינות המשתתפות בו, ובהן ישראל, לשם תרגום והתאמה. בישראל תורגמו המבחן והשאלונים לשתי שפות ההוראה העיקריות במדינה: עברית וערבית.

בשנת 2015 נערך מחקר חלוץ בקרב מדגם מצומצם של תלמידים בכיתה ד'. המחקר נועד לבדוק את התאמתם של כלי ההערכה ואת ההליכים והנהלים לאיסוף המידע. באמצעות הנתונים שנאספו בו נבדקה איכות התרגום לשפות השונות, ותוצאות מחקר החלוץ שימשו בסיס לבחירת הפריטים למחקר העיקרי. בשנת 2016 נערך המחקר העיקרי: המבחנים והשאלונים הועברו בקרב מדגם מייצג של תלמידים בכיתה ד' בכל אחת מהמדינות המשתתפות, וכן הועברו שאלונים למורים, למנהלים ולהורים של אותם תלמידים. לאחר מכן נבדקו המבחנים בכל מדינה, ונתוני המבחנים והשאלונים הועברו לעיבוד במרכז המחקר הבין-לאומי. המחקר הסתיים כשנה וחצי לאחר העברת המבחנים והשאלונים, עם פרסום תוצאותיו בדוח בין-לאומי מקיף בדצמבר 2017.¹

מחקר פירלס קובע סטנדרטים גבוהים וקפדניים לביצוע שלבי המחקר השונים, בהם שלב התרגום וההתאמה של המבחנים והשאלונים, דגימת בתי-הספר ודגימת הכיתות בתוכם, העברת המבחן בבתי-הספר ובדיקת המבחנים. הקפדה על סטנדרטים אלו מבטיחה את איכות הנתונים הנאספים ואת תקפותה של השוואת התוצאות בין המדינות המשתתפות בו. בפרק זה יתוארו השלבים העיקריים בביצוע המחקר בישראל. לפירוט על אודות שיטת המחקר וביצועו מנקודת מבט בין-לאומית של עורכי המחקר ראה המדריך הטכני של דוח פירלס 2016.²

3.1 כלי ההערכה במחקר פירלס

כלי ההערכה במחקר פירלס כוללים מבחן לתלמידי כיתה ד' הבודק את הישגיהם בקריאה של טקסטים מודפסים, וכן שאלונים לתלמידים הנבחרים, למורים שלהם לשפת אם, למנהלי בתי-הספר שבהם הם לומדים ולהורי התלמידים. השאלון לתלמיד משמש לאיסוף מידע על הרקע של התלמידים ועל עמדותיהם כלפי קריאה; השאלונים למורים ולמנהלי בתי-הספר משמשים לאיסוף מידע על ההקשר הלימודי של רכישת מיומנויות הקריאה ברמת הכיתה וברמת בית-הספר: רקע ועמדות של המורים, דרכי ההוראה, האקלים הלימודי בבית-הספר, ועוד. השאלון להורים משמש לאיסוף מידע על תמיכת הסביבה הביתית ברכישת הרגלי קריאה. מבחן אי-פירלס כולל מבחן לתלמידים הבודק את רמת הקריאה שלהם בסביבה ממוחשבת, וכן שאלון ייעודי קצר לתלמידים הנבחרים שהופיע בסוף המבחן וכלל שאלות על השימוש שלהם במחשב ובאינטרנט.

נוסף על כך קיים שאלון לנציג המדינה (בדרך כלל מתאם המחקר), המשמש לאיסוף מידע ברמה המערכתית על מדיניות הוראת המקצוע, על תכנית הלימודים בשפת אם בכלל ובקריאה בפרט ועל מדיניותה של מערכת החינוך בכל מדינה בנושאים קרובים. המידע שנאסף באמצעות השאלונים משמש לאפיון הזדמנויות הלמידה שניתנו לתלמידים ומספק את ההקשר של הישגיהם בקריאה ועמדותיהם כלפי קריאה. מידע על תכניות הלימודים ועל מדיניות מערכת החינוך בכל אחת מהמדינות המשתתפות מתפרסם בכרך שכותרתו "אנציקלופדיה של מחקר פירלס".³

¹ קישור לדוח הבין-לאומי של מחקר פירלס 2016: <https://timssandpirls.bc.edu/pirls2016/index.html>

² קישור למדריך הטכני של מחקר פירלס 2016:

<https://timssandpirls.bc.edu/publications/pirls/2016-methods.html>

³ קישור לאנציקלופדיה של מחקר פירלס 2016: <http://timssandpirls.bc.edu/pirls2016/encyclopedia/>

3.1.1 מבנה המבחן

בדומה למסגרת המושגית, גם שאלות המבחן ("פריטי המבחן") פותחו על ידי קבוצות מומחים להוראת קריאה במדינות שונות, בהתאם למערכת המושגים ולתאוריה של מסגרת המחקר (framework). הפריטים אושרו על ידי קבוצות מומחים בין-לאומיות שמונו על ידי מרכז המחקר הבין-לאומי.

כאמור בפרק 2, מסגרת המחקר מגדירה שתי מטרות עיקריות לקריאה: קריאה לצורך התנסות ספרותית וקריאה לצורך רכישת מידע. הפריטים במבחן פירלס (השאלות) מאורגנים ב-12 מקבצים מבוססי טקסטים (כל מקבץ כולל טקסט אחד). שישה מהטקסטים הם ספרותיים ונועדו לבדוק קריאה לצורך התנסות ספרותית, ושישה מידעיים ונועדו לבדוק קריאה לצורך רכישת מידע. בכל מקבץ 12 עד 17 פריטי מבחן הנלווים לטקסט ומתייחסים אליו (בעיקר בודקים את הבנתו). במבחן אי-פירלס הפריטים אורגנו בחמישה מקבצים מבוססי טקסט. כל הטקסטים היו מידעיים ונועדו לבדוק את מיומנויות הקריאה של התלמידים בסביבה דמוית אינטרנט.

הפריטים בודקים את מיומנויות הקריאה הפשוטות ואת מיומנויות הקריאה המורכבות (ראו פרק 2). כמחצית מהפריטים במבחן הם שאלות סגורות (שאלות רבות ברירה), וכמחציתם שאלות פתוחות ובהן התלמיד מתבקש להפיק תשובה ולבטאה בכתב. הטקסטים נבחרו כך שיתאימו לרמת הקריאות המצופה מתלמידים בשנתם הרביעית בבית-הספר ולעולם התוכן של ילדים בגיל זה במרבית התרבויות המקובלות במדינות המשתתפות. הטקסטים צריכים להיות ארוכים מספיק, לעורר עניין בתלמיד ולאתגר אותו, ובעיקר עליהם לזמן שאילת שאלות שיבדקו את המיומנויות שמבקשים למדוד במחקר, ולהתאים לשאילת שאלות סגורות ופתוחות כאחד.

בלוח 3.1 ו**בלוח 3.2** מוצגים החלק היחסי של מספר הפריטים במבחן בכל אחד מסוגי הטקסט ושל כל אחת ממיומנויות הקריאה הנבדקות, וכן מספר הפריטים מכל סוג טקסט ומיומנות קריאה. כמו כן, בעבור כל אחת ממיומנויות הקריאה הנבדקות נתון מספר הפריטים מסוג שאלות סגורות ומספר הפריטים מסוג שאלות פתוחות וחלקן היחסי בציון הכולל.

לוח 3.1: הפריטים במבחן - לפי סוג הטקסט וסוג פריט (רב ברירה - "סגור", או הפקת תשובה - "פתוח")

מבחן אי-פירלס				מבחן פירלס				סוג הטקסט
החלק היחסי במבחן	מספר הפריטים	פריטים סגורים	פריטים פתוחים	החלק היחסי במבחן	מספר הפריטים	פריטים סגורים	פריטים פתוחים	
0%	0	0	0	51%	90	46	44	טקסט ספרותי
100%	91	40	51	49%	85	40	45	טקסט מידעי
100%	91	40	51	100%	175	86	89	סה"כ

לוח 3.2: הפריטים במבחן - לפי מיומנויות קריאה וסוג פריט (רב ברירה - "סגור", או הפקת תשובה - "פתוח")

מבחן אי-פירלס				מבחן פירלס				מיומנויות הנבדקות
החלק היחסי במבחן	מספר הפריטים	פריטים סגורים	פריטים פתוחים	החלק היחסי במבחן	מספר הפריטים	פריטים סגורים	פריטים פתוחים	
21%	22	10	12	26%	50	25	25	אחזור מידע
28%	27	12	15	26%	53	35	18	הסקה ישירה
34%	23	7	16	35%	47	11	36	פרשנות ושילוב מידע
18%	19	11	8	13%	25	15	10	הערכת טקסט
100%	91	40	51	100%	175	86	89	סה"כ

במחקר פירלס הנוכחי, המבחן כלל לראשונה גם טקסטים פשוטים יותר, האמורים להתאים לתלמידים שטרם סיימו את רכישת הקריאה. טקסטים מסוג זה נכללו במחזור מחקר 2011 בתור מבחן נפרד (pre-PIRLS, פרה-פירלס), בו השתתפו מדינות עם הישגים נמוכים, להן הומלץ להשתתף במבחן זה, כדרך לשפר את התאמת הרמה של הטקסטים לשלב רכישת הקריאה של התלמידים. במחזור מחקר 2016 שאפו מארגני המחקר ליצור קישור בין פרה-פירלס לבין פירלס ולכן שילבו שני טקסטים מתוך מבחן פרה-פירלס במבחן פירלס, קרי שיעורם כשישית מכלל הטקסטים.

3.1.1.1 ארגון הפריטים בחוברות המבחן ובמבחן הממוחשב

במבחן פירלס, מתוך 12 המקבצים של המבחן, 10 מקבצים מאורגנים ב-15 גרסאות מבחן (חוברות). בכל גרסה יש שני מקבצים: מקבץ המבוסס על טקסט ספרותי ומקבץ המבוסס על טקסט מידעי, או שני מקבצים המבוססים על טקסט ספרותי או שניים המבוססים על טקסט מידעי. מטבע הדברים הייתה חפיפה חלקית בין גרסאות המבחן: כל אחד מהמקבצים הופיע בשלוש מבין 15 הגרסאות השונות. גרסה נוספת של המבחן כללה את שני המקבצים הנותרים, אחד מהם מבוסס על טקסט ספרותי ואחד על מידעי, אשר מוצגים לתלמיד בפורמט של מקראה, שבו הטקסטים והשאלות מופרדים לשתי חוברות שונות. הטקסטים כוללים תמונות צבעוניות, כדי לשפר את חווית הקריאה ואת האטרקטיביות של המבחן. כל תלמיד מקבל גרסה אחת מתוך 16 גרסאות המבחן.

במבחן אי-פירלס, 5 המקבצים של המבחן, מאורגנים ב-20 גרסאות מבחן. כל תלמיד נבחן בגרסה אחת, הכוללת שני מקבצים, כאמור במבחן זה כל הטקסטים מידעיים.

במערך כזה של מבחן, האופייני למבחנים שבהם מטרת המדידה היא המערכת בכללותה ולא התלמיד היחיד, כל תלמיד אמנם נבחן רק על שני מקבצי קריאה, אך במצטבר, כאשר מצרפים את תגובות כל התלמידים במדגם מייצג של מדינה נתונה, ניתן לקבל תמונה מקיפה ברמה ארצית על אודות השליטה של התלמידים (בשכבת הכיתה המוערכת) במדינה. מאידך, אי-אפשר (וכאמור, גם אין כוונה) לספק לתלמיד הנבחן ציון תקף, ואף לא לקבוצות קטנות כגון כיתה או בית-ספר, שכן בקבוצה המונה מספר יחסית מצומצם של תלמידים, מספר התגובות הנאספות על כל אחד מפריטי המבחן נמוך מידי, תלמידים שונים נבחנים בגרסאות מבחן שונות (ולכן ציוניהם אינם בני השוואה). זאת ועוד, התלמידים אינם בהכרח מהווים מדגם מייצג של הכיתה או שכבת הגיל בבית-הספר אליו משתייכים.

3.1.2 השאלונים

כמו המבחן, גם השאלונים במחקר פירלס פותחו על ידי קבוצות מומחים ממדינות שונות שמינה מרכז המחקר הבין-לאומי. בכל אחד מהשאלונים השאלות סגורות, ומרביתן מורכבות ממספר היגדים. ממלא השאלון מתבקש לקבוע את מידת הסכמתו עם ההיגד או את התדירות שבה הוא עושה פעולה מסוימת, על פני סולם בן ארבע או חמש דרגות. מקצת השאלות מבקשות לברר מידע מסוים (כגון השאלה "באיזו תדירות המורה שלך נותן לך שיעורי בית"?).

השאלון לתלמיד במבחן פירלס: התלמידים השיבו על השאלון העיקרי מיד בסיום מבחן פירלס. השאלון אחיד לכל התלמידים. הוא מכיל 20 שאלות סגורות ומתוכנן לכ-30 דקות השבה. השאלות עוסקות בהיבטים שונים בחיי התלמידים בבית ובבית-הספר, בהם מידע דמוגרפי בסיסי ומידע על אודות הסביבה הביתית, אקלים הלמידה בבית-ספר, הדימוי העצמי של התלמיד ועמדותיו כלפי קריאה. נושאי השאלון והמידע הנאסף בו עשויים להיות קשורים לרמת הישגים של התלמידים ולהסביר אותם.

השאלון לתלמיד במבחן אי-פירלס: בסיום מבחן אי-פירלס התלמידים השיבו על שאלון אחיד קצר שכלל 4 שאלות סגורות. השאלות עסקו במידת השימוש של התלמידים במחשבים, במידת השימוש באינטרנט לצורך קריאה וחיפוש מידע, וכן בגורם שבעזרתו רכשו התלמידים מיומנויות שונות הקשורות במחשבים (שימוש במחשב, איתור מידע, הקלדה).

השאלון למורים: בשאלון 32 שאלות סגורות העוסקות ברקע האקדמי והמקצועי של המורים, במשאבים בכיתה, בדרכי ההוראה ובעמדות המורים כלפי ההוראה. בשאלון גם נאסף מידע על מאפייני הכיתה הנבחרת, חומרי הלימוד, נושאי הלימוד, הפעילויות המשמשות ללימוד קריאה ולטיפוח העניין של התלמידים, ועוד.

השאלון למנהלי בתי-הספר: השאלון עוסק במאפייני בית-הספר. השאלון כולל 21 שאלות סגורות הנוגעות להיבטים שונים: מאפייני בית-הספר, משך הזמן המוקדש להוראת שפה, זמינות המשאבים והטכנולוגיה, מעורבות ההורים, אקלים הלמידה בבית-הספר, צוות המורים, תפקיד המנהל ומוכנות התלמידים.

השאלונים למורים ולמנהלים היו בפורמט אינטרנטי. קישור לשאלון נשלח למשיבים כשבועיים לפני מועד המבחן. משך הזמן המשוער המוקדש למילוי השאלון הוא כ-30-45 דקות, והמורים והמנהלים התבקשו לסיים למלאו לפני מועד העברת המבחנים והשאלונים לתלמידים.

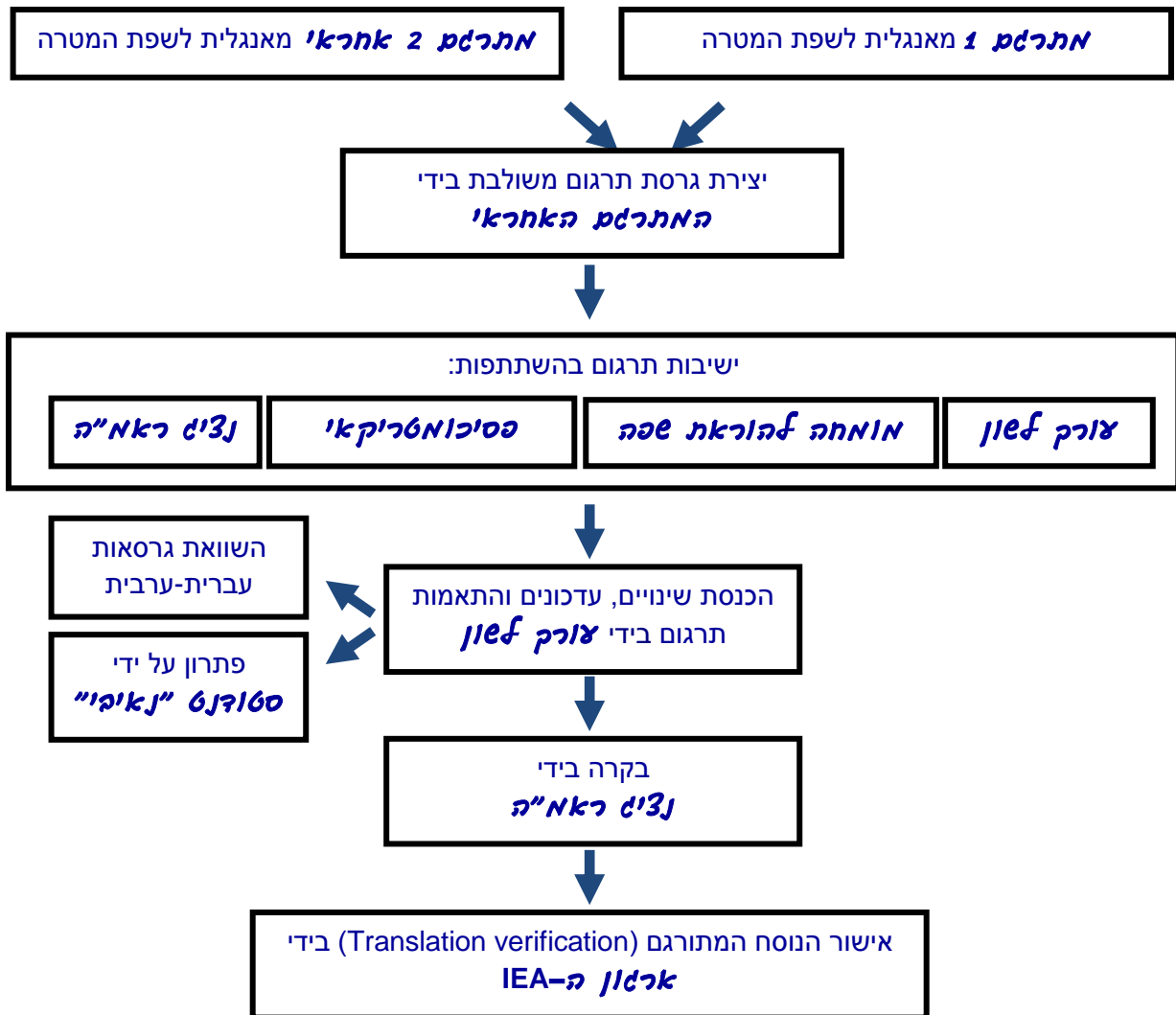
השאלון להורים: שאלון זה עוסק בתמיכת הסביבה הביתית ברכישת אוריינות הקריאה. השאלון כולל 21 שאלות סגורות ובהן שאלות רקע על השכלת ההורים ועיסוקם וכן על הרגלי הקריאה של ההורים ועמדותיהם כלפי קריאה, העיסוק בבית בפעילויות אוריינות מוקדמות כגון קריאת ספרים משותפת ולימוד קריאה בבית, והאם ילדיהם הלכו לגן. משך הזמן המשוער המוקדש למילוי השאלון הוא 15 עד 30 דקות.

3.1.3 תהליך התרגום וההתאמה של חומרי המחקר

כל חומרי המחקר בישראל, ובכלל זה פריטי המבחן והמחוננים לבדיקתם, השאלונים והמדריכים, תורגמו מאנגלית והותאמו בנפרד לכל אחת משתי שפות ההוראה בישראל, עברית וערבית. התרגום לשתי השפות נעשה במקביל, באמצעות צוותי תרגום ייעודיים וליווי ובקרה שוטפים של ראמ"ה (המנהלת את התהליך בכללותו). בסופו של התהליך נעשתה גם בקרה סופית על ידי מרכז המחקר הבין-לאומי. הליך התרגום של פריטי המבחן והשאלונים מוצג בפירוט ב**תפרשים 3.1**. תהליך מורכב זה מבטיח את איכות התרגום, את נאמנותו למקור מבחינת התוכן, את שימור המאפיינים הפסיכומטריים של הטקסטים ושל הפריטים (כגון משלב השפה, מורכבות המשפטים ואחידות המסיחים בשאלות הסגורות) ואת האחידות הסגונית בכל אחת משפות המטרה.

ארגון IEA מספק הנחיות ברורות בדבר ההתאמות שאפשר לעשות לפריטים. התאמה מקובלת היא, למשל, החלפת שמות הדמויות או המקומות בגרסת המקור בשמות מוכרים ומקובלים בשפת המטרה. בנוסף, הליך התרגום בישראל לווה בהליכים נוספים של בקרת איכות. כגון, שלב שבו מתקיים דיון של אנשי מקצוע בתחומים שונים לגבי כל אחד מהחומרים המתורגמים, שלב ההשוואה בין התרגום בעברית והתרגום בערבית ועוד. הליך התרגום המוקדף על שלביו השונים נועד להבטיח את איכות התרגום, את הנאמנות למקור, את שימור המאפיינים הפסיכומטריים של הטקסטים ושל הפריטים (כגון משלב השפה, מורכבות המשפטים ואחידות המסיחים בשאלות סגורות), ואת האחידות הסגונית של יחידות המבחן, וזאת משום שלהיבטים האלו עשויה להיות השפעה על רמת הקושי של הטקסטים ושל הפריטים השואלים אודותיהם.

תרשים 3.1: תהליך התרגום בישראל משפת המקור (אנגלית) לכל אחת משפות המטרה (עברית/ערבית)



באשר למבחן אי-פירלס, בישראל קיים גם צורך להתאים את כל רכיבי המבחן הממוחשב לשפות הנכתבות מימין לשמאל. כאשר מתבצע הליך של תרגום והתאמת מבחן שתוכן בשפה הנקראת משמאל לימין (אנגלית) לשפה הנקראת מימין לשמאל (עברית וערבית בישראל), יש לוודא כי הכיוון של הרכיבים השונים המוצגים בפריטי המבחן (רכיבי טקסט וייצוגים גרפיים אחרים) וכן כיוון הממשק של מערכת המבחן מותאמים כולם לכיווני הקריאה בשפה. כיוון קריאת הטקסט בשפה מכתוב את רצף העיבוד הקוגניטיבי של המידע המוצג בפריט, דבר המשליך על המראה של הרכיבים עצמם, על מיקומם בפריט ועל הסידור שלהם במסך. ככלל, כל רכיב לא טקסטואלי בממשק יעבור התאמה לכיוון הקריאה והמיקום שלו יתהפך כבתמונת ראי. לדוגמה, בגרסת המקור הטקסט הופיע בצד השמאלי של המסך והשאלות מימין, בגרסאות המבחן בישראל הטקסט הופיע בצד הימני של המסך והשאלות משמאלו.

לאחר סיומו של הליך התרגום, הועברו החומרים המתורגמים למרכז המחקר הבינלאומי לשם אימות ובקרה על התרגום ואישור סופי שלהם על ידי מארגני המחקר.

3.2 מחקר החלוץ

כל פריטי המבחן שפותחו לצורך המחקר הנוכחי, וכן שאלוני הרקע, נוסו במחקר חלוץ על מדגם מצומצם של תלמידים ובתי-ספר. המחקר נערך במהלך שנת 2015, כשנה לפני המחקר העיקרי, בכל אחת מהמדינות המשתתפות. מחקר החלוץ בישראל נערך בחודשים מרץ-אפריל 2015, בקרב 26 בתי-ספר שנדגמו באופן אקראי על-ידי הארגון הבינלאומי. בכל בית-ספר נדגמו באופן אקראי שתי כיתות אם מבין כיתות ד' בבית-הספר. תלמידי שתי הכיתות השתתפו במחקר פירלס ומילאו במסגרתו מבחן ושאלון. ובנוסף השתתפו גם המורים לשפה והמנהלים באותם בתי-ספר. אחת מתוך שתי הכיתות בכל אחד מ-26 בתי-הספר, השתתפה גם במבחן אי-פירלס כחודש לאחר העברת מחקר פירלס בבית-הספר.

למחקר החלוץ שתי מטרות: לבדוק את הפריטים ולתרגל את הנהלים הלוגיסטיים הכרוכים בביצוע המחקר. בהיבט הבין-לאומי, בדיקת הפריטים משמשת לבחירת המקבצים (טקסטים ופריטים) שייכללו במחקר העיקרי ולאיתור השינויים הנדרשים בפריטים עד לגיבושם הסופי. בהיבט הלאומי, בדיקת הפריטים מאפשרת לאמת את טיב התרגום ואת ההתאמה הן בתוכן והן ברמת המשלב והבהירות של הפריטים.

כדי להבטיח באופן מיטבי את השגת מטרות מחקר החלוץ, חולק המדגם כך שמחצית מבתי-הספר שנדגמו היו דוברי עברית ומחציתם דוברי ערבית, וזאת בשונה מהפרופורציה של שתי קבוצות אלו באוכלוסייה. חלוקה זו אפשרה לקבל נתונים איכותיים אודות תפקוד הטקסטים והפריטים בכל אחת מהשפות, עברית וערבית, היות שמספר גדול מספיק של תלמידים השיב על כל אחד מהפריטים.

בעקבות מחקר החלוץ סיפק המרכז הבין-לאומי לכל מדינה דוח נתונים פסיכומטריים של פריטי המבחן, כאשר בעבור ישראל הופקו דוחות נפרדים לפריטים בעברית ולפריטים בערבית. נתונים אלו אפשרו לבחון את תפקודם של הפריטים ברמה הלאומית ולאתר קשיים ופריטים חריגים בנוסח המתורגם בכל מדינה שהשתתפה במחקר ואשר שפתה אינה אנגלית. מנהלי המחקר בכל מדינה נתבקשו לשוב ולבדוק פריטים שנתוניהם נמצאו חריגים ולברר אם אין בהם בעיות בתרגום או בהפקה (כגון בעיות דפוס, עיצוב גרפי וכיוצא בזה).

נתוני מחקר החלוץ עובדו ונותחו במרכז המחקר הבין-לאומי, ועל פיהם בחר המרכז את הטקסטים ואת הפריטים שייכללו במחקר העיקרי. במחקר פירלס, מתוך 12 הטקסטים שנכללו במחקר החלוץ, נבחרו 6 טקסטים שישמשו במחקר העיקרי, לצד טקסטים ממחזורי מחקר פירלס קודמים. בנוסף, חלק מהפריטים עברו שינוי בעקבות ניתוחי הפריטים. השינויים נעשו על פי שיקולים פסיכומטריים (למשל: נגרעו מהמחקר פריטים שנמצאו קשים מדי או קלים מדי ברמה הבין-לאומית, או פריטים שאינם מבחינים במידה מספקת בין תלמידים בעלי רמות מיומנויות קריאה שונות). על סמך מחקר החלוץ גובשו חומרי ההערכה הסופיים: המבחן וכן השאלונים לתלמיד, למורה ולמנהל. במבחן אי-פירלס, מתוך 6 הטקסטים שנכללו במחקר החלוץ, הוסר טקסט אחד במעבר למחקר העיקרי וכן בוצעו שינויים בפריטים משיקולים פסיכומטריים.

3.3 המחקר העיקרי

השלב העיקרי במחקר – שלב העברת המבחנים והשאלונים בבתי-הספר – נערך בישראל בחודשים מאי-יוני 2016. להלן יתוארו השלבים המרכזיים בביצוע המחקר העיקרי.

3.3.1 הדגימה

במחקר העיקרי נדרשים להשתתף לפחות 150 בתי-ספר ולפחות 4,500 תלמידים מכל אחת מהמדינות המשתתפות. שיטת הדגימה, המחייבת מדינות אלו, נקבעת על ידי מרכז המחקר הבין-לאומי ונועדה להבטיח כי מדגם זה של בתי-ספר ושל תלמידים אכן מייצג את אוכלוסיית המטרה באותה מדינה. זאת ועוד: הדגימה בפועל נעשית אף היא על ידי מרכז המחקר הבין-לאומי, על סמך

הנתונים המועברים אליו מהמדינות המשתתפות. בשל השונות הגדולה הקיימת בהישגי התלמידים בישראל, כפי שבאה לידי ביטוי במחזורי מחקר פירלס קודמים וכן במחקרים בינלאומיים אחרים, המליץ הארגון הבינלאומי לישראל להגדיל מעט את מספר בתי-הספר הנדגמים לצורך המחקר העיקרי. בישראל נדגמו לצורך ההשתתפות במחקר 160 בתי-ספר.

על פי מדיניות מארגני מחקר פירלס 2016, נקבע כי על המדינות המבקשות להשתתף במבחן אי-פירלס חלה חובה להשתתף גם במבחן פירלס (אך לא להיפך). במדינות שהשתתפו במחקר פירלס 2016 ושברו להשתתף בשני המבחנים, בוצעה דגימה אחת משותפת בעבור שני המבחנים, כך שאותם התלמידים אמורים היו להשתתף גם בפירלס וגם באי-פירלס.⁴ הדבר סייע בידי עורכי המחקר לכייל את נתוני הישגי התלמידים בשני המבחנים לאותו סולם.

3.3.1.1 מערך הדגימה

מערך הדגימה ששימש במחקר פירלס 2016 (כמו במחזורי קודמים של המחקר) הוא מערך דו-שלבי בשכבות (Stratified Two-Stage Cluster Sample Design): בשלב הראשון נדגמים באקראי בתי-ספר. בשלב השני נדגמת באקראי כיתה אחת בכל בית-ספר שנדגם. כל התלמידים בכיתה שנדגמה אמורים להשתתף (למעט מספר קטן מאוד של תלמידים, בהתאם לכללי המחקר המפורטים בתיבה 3.2).

כדי להבטיח ייצוג הולם במדגם לכל הקבוצות והמגזרים באוכלוסייה (מגדר, מגזר שפה,⁵ בתי-ספר מסוגים שונים, רקע חברתי-כלכלי וכיוצא בזה), מקבצים בהליך הכנתו את בתי-הספר לשכבות (strata), ומהן יידגמו בתי-הספר על פי שיעורה של כל שכבה באוכלוסייה. שכבות הדגימה מוגדרות בכל מדינה בשיתוף מרכז המחקר הבין-לאומי. בישראל תוכננו שכבות הדגימה כך שייצגו היבטים מרכזיים במערכת החינוך בארץ וישקפו את השונות הקיימת במוסדות החינוך השונים, ותובא בחשבון השפעתם האפשרית והקשרים השונים בין המשתנים השונים להישגי התלמידים (למשל: הקשר בין הישגים לימודיים למשתנים "מגזר שפה", "רקע חברתי-כלכלי" וכיוצא באלה). השכבות השונות (הגלויות והסמויות) ששימשו במחקר פירלס 2016 בישראל מוצגות בלוח 3.3.

דגימת בתי-הספר נעשית ביחס לגודלו - Systematic Probability Proportional to Size (PPS). ליתר פירוט, בשיטה זו, הסיכוי של בית-ספר להידגם למחקר גבוה יותר ככל שבית הספר גדול יותר. לפיכך, אם למשל בית-ספר אחד גדול פי שניים מבית-ספר אחר, ההסתברות שהוא יעלה במדגם תהיה כפולה מזו של בית הספר הקטן. בתוך בית-ספר שנדגם, הסיכוי של כיתה להידגם למחקר נמוך יותר בבית הספר הגדול מאשר בבית הספר הקטן, ביחס הפוך למספר הכיתות בשכבה ה'. השילוב בין דגימה פרופורציונלית של בית-הספר לפי גודלו ודגימה תוך בית-ספרית מבטיח שהסיכויים של כל תלמיד ותלמיד באוכלוסייה להידגם למחקר דומים, ואינם תלויים בגודל בית-הספר שבו הוא לומד. בכך נמנע מצב שתלמידים הלומדים בבתי-ספר קטנים יידגמו בהסתברות גבוהה יותר מתלמידים הלומדים בבתי-ספר גדולים.

כלל בית-ספר שנדגם למחקר נדגמים מלכתחילה שני בתי-ספר "מחליפים" (מחליף ראשון ומחליף שני), הדומים במאפייני הרקע שלהם למאפייניו של אותו בית-ספר. ישנן לדוגמה סיבות טכניות להחלפת בתי-הספר, כגון סגירת בית-ספר או איחוד עם בית-ספר אחר.

⁴ חריגה לכך היא קנדה: מבין 8 הפרובינציות שהשתתפו במחקר פירלס, רק 4 בחרו להשתתף במבחן אי-פירלס, ועל כן הכיסוי של אוכלוסיית המטרה במבחן אי-פירלס עמד על 74% בלבד (לעומת 97% במבחן פירלס).

⁵ "מגזר שפה" מתייחס לחלוקת בתי-הספר בהתאם לשפת הדיבור השגורה בהם: בתי-ספר שרוב תלמידיהם דוברי עברית נקראים ב"דוח זה "בתי-ספר דוברי עברית", ובתי-ספר שרוב תלמידיהם דוברי ערבית נקראים "בתי-ספר דוברי ערבית".

לוח 3.3: שכבות הדגימה במחקר פירלס 2016 בישראל

שכבת דגימה-1:	שכבת דגימה-2:	שכבת דגימה-3:
מגזר שפה וסוג פיקוח	רקע חברתי-כלכלי / אזור גאוגרפי*	מגזר בבתי"ס שבפיקוח ממ"ד
בתי-ספר דוברי פיקוח ממלכתי	נמוך, בינוני, גבוה	
ערבית פיקוח ממלכתי-דתי	נמוך, בינוני, גבוה	בנים, בנות, מעורב
בתי-ספר דוברי ערבים	נמוך, בינוני	
ערבית דרוזים		
ערבית בדואים	צפון, דרום	

לוח 3.3 אפשר לראות כי המשתנים שלפיהם גובשו שכבות הדגימה בישראל הם מגזר שפה, סוג פיקוח, לאום/מגזר (ערבים, דרוזים, בדואים), ורקע חברתי-כלכלי של בית-הספר (להרחבה ראו **תיבה 3.1**) או שיוך לאזור גאוגרפי (צפון לעומת דרום – רק בקרב בתי-ספר במגזר הבדואי). תת-חלוקה נוספת בבתי-ספר דוברי עברית שבפיקוח הממלכתי-דתי היא על פי הפרדה מגדרית (בתי-ספר לבנים, לבנות, או מעורבים).

תיבה 3.1: מדד הטיפוח החברתי-כלכלי "שטראוס" והשימוש בו במחקר זה

מדד הטיפוח "שטראוס" הוא מדד אישי המחושב לכל תלמיד ותלמיד בבית-ספר (בכל שכבות הכיתה הקיימות בו). זהו מדד המשקף את הרקע החברתי-כלכלי של התלמיד, והוא נהוג במערכת החינוך בישראל החל ממצחית העשור הראשון של המאה ה-21. משוקללים בו ארבעה רכיבים: (א) השכלת ההורים – מבוסס על רמת ההשכלה של ההורה המשכיל יותר (חלקו במדד 40%); (ב) מצב כלכלי – מבוסס על חציון הכנסת הורי התלמידים בבית הספר (חלקו במדד 20%); (ג) פריפריאליות – מבוסס על המרחק הפיזי בין מיקומו של בית הספר לריכוזי אוכלוסייה גדולים (חלקו במדד 20%); (ד) מדד המתחשב בשילוב של הגירה (של התלמיד עצמו או לפחות של אחד מהוריו) וארץ מצוקה - מבוסס על רשימת ארצות מצוקה של האו"ם (חלקו במדד 20%).

מדד הטיפוח "שטראוס" מבוסס באמצעות עשירוני טיפוח המייצגים את עשירון התלמידים בהתפלגות המדד בכלל התלמידים בבתי-הספר בארץ (מעשירון 1- המבוססים ביותר, ועד לעשירון 10- החלשים ביותר). על בסיס עשירוני הטיפוח סווגו התלמידים לשלוש קבוצות של רקע חברתי-כלכלי אישי: גבוה (עשירונים 1-3), בינוני (4-7) ונמוך (8-10) לצורך הניתוח והדיווח במחקר זה.

הרקע החברתי-כלכלי של בית הספר נקבע כממוצע הרקע החברתי-כלכלי האישי של כלל התלמידים בבית"ס.

עוד על מאפייני מדד הטיפוח ראו בקישור:

<http://cms.education.gov.il/EducationCMS/Units/Scientist/CareIndex/>

3.3.1.2 אוכלוסיית המטרה, מסגרת הדגימה והמדגם

מחקר פירלס מגדיר את **אוכלוסיית המטרה** בכל מדינה כך: "כל התלמידים הלומדים בדרגת כיתה המייצגת את שנת הלימודים הרביעית בבית-הספר, אלא אם כן הגיל הממוצע של תלמידים בשנת הלימודים הרביעית נמוך מ-9.5 שנים"⁶. מתוך אוכלוסיית המטרה בכל מדינה נבנית **מסגרת הדגימה**, וממנה נדגמים בתי-הספר והכיתות שישתתפו במחקר. על פי הגדרה זו, ברוב המדינות, וגם בישראל, אוכלוסיית המטרה היא כלל התלמידים הלומדים בכיתות ד' (אוכלוסיית מטרה זו

⁶ קביעת השנה הראשונה בבית-הספר בכל מדינה נעשית על פי תקן בין-לאומי לסיווג רמות לימוד במדינות שונות ה-ISCED International Standard Classification of Education (ISCED), שפותח על ידי אונסקו. ה-ISCED מספק סכמה לסיווג רמות הלימוד החל בטרום-יסודי (רמה 0) וכלה בתואר דוקטור (רמה 6 לפי ISCED 1997 ורמה 8 לפי ISCED 2011).

מוגדרת "100% התלמידים").⁷ מתוך אוכלוסיית המטרה מוגדרת "אוכלוסיית התלמידים שאמורים להיבחן". הסטנדרטים של מחקר פירלס מאפשרים למדינות שלא לכלול באוכלוסיית התלמידים האמורים להיבחן עד 5% מהתלמידים שבאוכלוסיית המטרה, וזאת באמצעות גריעתם של בתי-ספר שלמים מתוך אוכלוסיית המטרה או גריעתן של כיתות שלמות בתוך בתי-ספר או אף של תלמידים יחידים הלומדים בכיתות שנדגמו (לפירוט הכללים לגריעת בתי-ספר, כיתות או תלמידים כפי שהוגדרו במחקר ראו **תיבה 3.2**). בלוחות המציגים את תוצאות המחקר בדוח הבין-לאומי נרשמת הערה ליד שמות המדינות שבהן אושר שיעור גריעה גבוה מהסטנדרט האמור. במחקר פירלס 2016, כמו במחזורי מחקר קודמים, אישר מרכז המחקר הבין-לאומי לישראל לגרוע מאוכלוסיית המטרה תלמידים בשיעור גבוה מ-5%, וזאת בשל ייחודיותה של מערכת החינוך בישראל על זרמיה וגוניה, ובעיקר בשל אי-השתתפותם של בתי-הספר החרדיים במחקרים הבין-לאומיים בעבר ובהווה.

תיבה 3.2: כללי מחקר פירלס לגריעת תלמידים מאוכלוסיית המטרה

הכללים המוגדרים במחקר פירלס לגריעת תלמידים מאוכלוסיית המטרה הם אלה: **בשלב דגימת בתי-הספר** אפשר לגרוע (באמצעות הוצאה ממסגרת הדגימה) בתי-ספר "שאינם נגישים" (בשל ריחוק גיאוגרפי, למשל), בתי-ספר "קטנים מאוד" (שבהם השכבה הרלוונטית למחקר מונה 4 תלמידים או פחות), בתי-ספר שהמבנה או תכניות הלימודים שלהם שונים מהמבנה או מתכניות הלימודים של בתי-הספר המשתייכים לזרם העיקרי במדינה (למשל, בישראל - בתי-ספר חרדיים), ובתי-ספר שכלל הלומדים בהם פטורים מראש מלהשתתף בבחינה (למשל, בתי-ספר לחינוך מיוחד). **בשלב דגימת הכיתות** אפשר לגרוע כיתות של תלמידי חינוך מיוחד בבתי-הספר שנדגמו. **לאחר שנדגמו הכיתות**, אפשר לגרוע תלמידים יחידים מכיתה שנדגמה ולפטור אותם מלהשתתף במחקר על רקע של צרכים מיוחדים כגון לקויות למידה ו/או נכות פיזית שאינה מאפשרת להם להשתתף במבחן, או על רקע שפתי (תלמידים מהגרים, הלומדים פחות משנה במדינה שאליה היגרו, ולכן עדיין אינם שולטים בקריאה בשפת המבחן). יש לציין כי מרכז המחקר הבין-לאומי מאפשר לגרוע תלמידים אלו, בין היתר בשל העובדה שלא מתאפשרות חלופות הולמות כגון מתן התאמות מיוחדות ומתן תנאים מותאמים במבחן.

3.3.1.3 אוכלוסיית המחקר ומסגרת הדגימה בישראל

בישראל כללה אוכלוסיית המטרה בסך הכול 139,788 תלמידים שלמדו ב-2,572 בתי-ספר.⁸ בשלב דגימת בתי-הספר נגרעו **ממסגרת דגימה זו** בתי-ספר של החינוך המיוחד, בתי-ספר קטנים במיוחד ובתי-ספר ששפת ההוראה בהם היא אנגלית או צרפתית, בסה"כ כ-1.1% מהתלמידים. כמו כן נגרעו ממסגרת הדגימה בתי-ספר של הפיקוח החרדי (19.9% מכלל התלמידים באוכלוסיית המטרה), זאת בשל אי-השתתפותו של הזרם החרדי במחקר פירלס (בתי-ספר של הפיקוח החרדי נגרעו גם במחזורי מחקר פירלס קודמים). בסך הכול, שיעור הגריעה של תלמידים מאוכלוסיית המטרה בישראל עקב גריעת בתי-ספר ממסגרת הדגימה עמד על כ-21%. לאחר גריעת בתי-ספר אלו נותרו **במסגרת הדגימה** 1,696 בתי-ספר ובהם 110,408 תלמידים. מתוך בתי-הספר הללו נדגמו לצורך השתתפות במחקר 160 בתי-ספר.

טרם דגימת הכיתות שישתתפו במחקר, נגרעו מכל בתי-הספר שעלו במדגם כיתות של החינוך המיוחד. לאחר דגימת הכיתות, ניתן פטור לתלמידים יחידים העונים לקריטריונים המקנים פטור

⁷ במחקר זה, כמו במחזורי קודמים של מחקר פירלס בישראל, בתי-ספר דוברי הערבית במזרח ירושלים לא נכללים מראש בהגדרה של אוכלוסיית המטרה, זאת בעיקר משום שבבתי-ספר אלו לומדים, מסיבות היסטוריות, לפי תכנית הלימודים של ירדן.

⁸ הנתונים נכונים לחודש דצמבר 2014.

מהשתתפות במחקר (כמפורט בתיבה 3.2).⁹ בסך הכול עמד שיעור הגריעה של תלמידים מאוכלוסיית המטרה בישראל על כ-25% (מתוכם, כאמור, החלק הארי הם תלמידי המגזר החרדי).

3.3.1.4 ייצוגיות המדגם בפועל

שיעור הגריעה בישראל אמנם גבוה בהשוואה לשיעורה בשאר המדינות המשתתפות, אך מנגד נרשם בה שיעור השתתפות גבוה. שיעור השתתפותם של בתי-הספר שנדגמו בישראל היה מעל 99%. מתוך 160 בתי-הספר שנדגמו, שני בתי-ספר סירבו להשתתף במחקר, והוחלפו בבתי-ספר אחרים הדומים להם במאפייני הרקע, ואילו בית-ספר אחד נגרע על-ידי מארגני המחקר בשל שיעור השתתפות נמוך.¹⁰

שיעור השתתפותם של תלמידים בכיתות שנדגמו עמד על 95% במבחן פירלס ועל 91% במבחן אי-פירלס.¹¹ **בלוח 3.4** מוצגים התפלגות התלמידים לפי מגזר שפה ולפי סוג פיקוח בבתי-הספר המיוצגים במסגרת הדגימה, ושיעור המשתתפים בפועל במבחן פירלס.

לוח 3.4: תיאור המשתתפים במסגרת הדגימה ובמחקר בפועל במבחן פירלס

מגזר שפה	סוג פיקוח	מסגרת הדגימה			המדגם בבתי-הספר שהשתתפו			המשתתפים בפועל		
		מספר בתי"ס	מספר התלמידים	שיעור התלמידים	מספר בתי"ס	מספר התלמידים	שיעור התלמידים	מספר בתי"ס	מספר התלמידים	שיעור התלמידים
ממלכתי		860	58,918	53%	77	2,228	51%	77	2,063	51%
דוברי עברית	ממלכתי-דתי	410	20,347	19%	34	854	20%	34	785	19%
	סך הכול	1270	79,265	72%	111	3,082	71%	111	2,848	70%
דוברי ערבית	סך הכול	426	31,143	28%	48	1,286	29%	48	1,193	30%
ישראל כלי	סך הכול	1,696	110,408	100%	159	4,368	100%	159	4,041	100%

⁹ יש לציין כי בישראל, שיעור התלמידים המאובחנים כ'תלמידים בעלי צרכים מיוחדים' גבוה משיעורם בכלל המדינות המשתתפות במחקר. כדי לצמצם במידת האפשר את שיעור התלמידים הפטורים מלהשתתף במחקר בתוך כל כיתה שנדגמה, הנחו עורכי המחקר בישראל את בתי-הספר לבחור לא יותר מתלמיד אחד בעל צרכים מיוחדים בכל כיתה שיקבל פטור מהשתתפות במחקר. במקרים חריגים אושר פטור מהשתתפות במחקר ליותר מתלמיד אחד בכיתה, אך עורכי המחקר דאגו לאזן זאת באמצעות כיתות אחרות בהן לא נתקבלה בקשה לפטור תלמידים.

¹⁰ בית-ספר אחד אמנם השתתף במחקר, אך נגרע ונחשב על ידי מארגני המחקר כאילו שלא השתתף, שכן שיעור ההשתתפות של התלמידים בו היה נמוך מ-50%. לכן, בנתוני מבחן פירלס הבין-לאומי והישראלי נכללים תלמידים מ-159 בתי-ספר. כמו כן, שני בתי-ספר המשתתפים לפיקוח הממ"ד סרבו להשתתף במבחן אי-פירלס בלבד, ולכן בנתוני מבחן אי-פירלס נכללים תלמידים מ-157 בתי-ספר.

¹¹ ליתר פירוט: שיעור ההשתתפות של התלמידים במחקר מחושב לפי מספר התלמידים שהשתתפו במחקר בפועל מתוך כלל התלמידים שהיו אמורים להשתתף בו (תלמידים בכיתות המדגם בניקוי התלמידים הפטורים מהשתתפות, כמפורט בתיבה 3.2) בבתי-ספר שהשתתפו במחקר (קרי, לא כולל בתי-ספר שסרבו להשתתף ולא הוחלפו, או בתי-ספר בהם שיעור ההשתתפות היה נמוך מערך 50%). בישראל, מספר התלמידים שהיו רשומים כמי שלומדים בכיתות שנדגמו למחקר, בבתי-הספר שהשתתפו במבחן פירלס, היה 4,368. מתוכם 13 תלמידים, כך הסתבר, עזבו את בית-הספר אף שעדיין היו רשומים בכיתות שנדגמו, ו-107 תלמידים קבלו פטור ממרכז המחקר בישראל, על רקע של צרכים מיוחדים או רקע שפתי. לאחר גריעתם של אלו, נותרו 4,248 תלמידים שהיו אמורים להשתתף במחקר. מתוכם נעדרו 207 תלמידים מבית-הספר ביום המבחן. סה"כ השתתפו במבחן פירלס בישראל 4,041 תלמידים, שהם כ-95% מהתלמידים שהיו אמורים להשתתף בו.

מספר התלמידים שהיו רשומים כמי שלומדים בכיתות שנדגמו למחקר, בבתי-הספר שהשתתפו במבחן אי-פירלס, היה 4,315. מתוכם 14 תלמידים, כך הסתבר, עזבו את בית-הספר אף שעדיין היו רשומים בכיתות שנדגמו, ו-105 תלמידים קבלו פטור ממרכז המחקר בישראל, על רקע של צרכים מיוחדים או רקע שפתי. לאחר גריעתם של אלו, נותרו 4,196 תלמידים שהיו אמורים להשתתף במבחן אי-פירלס. מתוכם נעדרו 398 תלמידים מבית-הספר ביום המבחן (של מבחן אי-פירלס או פירלס), או שלא יכלו להשלים את המבחן בשל תקלה טכנית. סה"כ השתתפו במבחן אי-פירלס (ובמבחן פירלס) בישראל 3,798 תלמידים, שהם כ-91% מהתלמידים שהיו אמורים להשתתף בו.

מלוח 3.4 עולה כי שיעור התלמידים שהשתתפו בפועל במחקר, בשני מגזרי השפה, דומה לשיעורם באוכלוסייה (כמוגדר במסגרת הדגימה), אם כי לא זהה. חשוב לשים לב כי כדי "לפצות" על ההבדלים בפועל בשיעורי הדגימה וההשתתפות בכל שכבה לעומת מסגרת הדגימה, נעשה על ידי מארגני המחקר בעבור כל מדינה ומדינה המשתתפת במחקר "משקול" של תוצאות התלמידים במבחן בשכבות השונות כך שתוצאות המחקר ייצגו כהלכה את מסגרת הדגימה (על משקול הנתונים ראה **תיבה 3.3**).

כאמור לעיל, כל התלמידים שנדגמו למחקר פירלס אמורים היו להשתתף בשני המבחנים, גם פירלס וגם אי-פירלס. בישראל, כמו בכל אחת מן המדינות המשתתפות במחקר, בנוסף להיעדרות מאחד המבחנים (לדוגמה, מפאת מחלה) הרי שתלמידים לעיתים לא יכלו להשלים את מבחן אי-פירלס בשל תקלות טכניות הנוגעות למחשבים. בישראל, 4,041 תלמידים השתתפו במבחן פירלס (הישגיהם מדווחים בפרק 4) ו-3,933 תלמידים השתתפו במבחן אי-פירלס, כאשר 3,798 השתתפו בשני המבחנים (הישגיהם מדווחים בפרק 5).¹²

תיבה 3.3: משקול הנתונים - כיצד חושב הציון הממוצע?

המדגם במחקר פירלס נעשה באמצעות שכבות דגימה שתוכננו מראש כך שהמדגם ייצג נאמנה את כלל האוכלוסייה על הקבוצות השונות המרכיבות אותה. בפועל, במחקר עצמו, בשל שיעורי ההשתתפות שונים בשכבות השונות, שיעור המשתתפים עלול שלא לייצג נאמנה את שיעור התלמידים בשכבות השונות באוכלוסייה. כדי "לפצות" על הבדלים אלו, לכל תלמיד שהשתתף במחקר מחושבת "משקולת" (weight), הנותנת משקל שונה לנתוני התלמיד בעת חישוב הממוצע הארצי (או המגזרי) ומדדים סטטיסטיים נוספים ברמה הארצית. המשקולות חושבו לפי משתני השכבות 1 ו-2 המוצגים לעיל ב**לוח 3.3** (מגזר שפה, סוג פיקוח ורקע חברתי-כלכלי של בית הספר), כך שלאחר קביעתן יהיו השיעורים המתוקנים בכל שכבה דומים לשיעורי השכבה המתאימה באוכלוסייה (למעשה, במסגרת הדגימה).

3.3.1.5 המורים, המנהלים וההורים שהשתתפו במחקר

במחקר פירלס מילאו שאלונים גם המורים המלמדים שפה בכיתות שנדגמו להשתתף במחקר, מנהלי בתי-הספר והורי התלמידים המשתתפים, והם מקושרים להישגי התלמידים במבחן ולשאלוני התלמידים. חשוב לציין כי המורים שמילאו את השאלונים אינם מדגם פשוט של מורים לשפה בכיתות ד', והנתונים בדוח זה מנותחים על פי שיעור התלמידים שלמדו אצל מורים בעלי מאפיין נתון, ולא על פי שיעור המורים בעלי המאפיין הזה. גם הנתונים משאלון ההורים מנותחים ברמת התלמידים. בכל בית-ספר שהשתתף במחקר, מנהל בית-הספר, המורה להוראת שפת אם של הכיתה שנדגמה והורי התלמידים שהשתתפו במחקר התבקשו למלא שאלוני ייעודיים. כמפורט ב**לוח 3.4** ו**בהערה 29**, במחקר השתתפו 159 מנהלים ומספר דומה של מורים להוראת שפת אם. כמו כן, כ-78% מהורי התלמידים שהשתתפו במחקר מילאו את השאלונים.

3.3.2 הליך איסוף הנתונים

כחודשיים לפני ביצוע המחקר, קיבלו מנהליהם של בתי-הספר שנדגמו הודעה על כך, הן בשיחה טלפונית אישית מנציגי ראמ"ה (מרכז המחקר בישראל) והן במכתב מטעם ראמ"ה. מנהלי בתי-הספר נתבקשו למנות איש קשר מטעמם, לרוב מורה מנוסה מהצוות החינוכי בבית-הספר או דמות ניהולית (כגון רכז שכבה, רכז פדגוגי, סגן מנהל). איש הקשר היה מופקד על הארגון הלוגיסטי של עריכת המחקר בבית-ספרו, לרבות תיאום מועד לקיומו, הקצאת חדר מתאים, איסוף שמות התלמידים בעלי צרכים מיוחדים, הודעה לתלמידים שנדגמו על השתתפותם במחקר ועידודם להשתתף בו ברצינות

¹² ביתר פירוט: 243 תלמידים השתתפו רק במבחן פירלס ו-135 תלמידים השתתפו רק במבחן אי-פירלס.

הראויה, וכיצא באלה פעולות לתיאום ביצועו של המבחן. כחודש לפני תחילת המחקר בישראל יזמה ראמ"ה מפגש היערכות למנהלים ולאנשי הקשר בבתי-הספר שנדגמו. במפגש הוצג לפנייהם מחקר פירלס – ובכלל זה הן מבחן פירלס להערכת קריאה של טקסטים מודפסים והן מבחן אי-פירלס להערכת קריאה בסביבה ממוחשבת – חשיבותו וחשיבות תוצריו, וניתנו הנחיות מפורטות להיערכות בית-הספר למחקר וכן דגשים להעלאת המוטיבציה של התלמידים להשתתף בו, לעשות במבחנים כמיטב יכולתם ולמלא את השאלונים ברצינות הראויה. הודגש בפני המנהלים ואנשי הקשר כי אין לעשות כל הכנה לימודית מיוחדת עם התלמידים בשכבת כיתות ד' בכלל ועם הכיתה שנדגמה להשתתף במחקר בפרט, וכי הכנה שכזו עשויה לחבל בתוקף המחקר ובמידה שבה המדגם מייצג את האוכלוסייה. בסיום המפגש נמסר לכל אחד מהנוכחים איזו כיתה מתוך כיתות ד' בבית-הספר נדגמה להשתתף במחקר.

השאלונים למנהלים ולמורים המלמדים שפת אם בכיתה שהשתתפה במחקר הועברו באמצעות האינטרנט. קישור אישי לשאלונים נשלח אליהם במייל כחודש לפני המועד שנקבע להעברת המבחן בבית-הספר, והם התבקשו להשלים את מילויים עד כשבוע לפני מועד העברת מבחן פירלס בבית-הספר. כל אחד מהמנהלים והמורים זוהה באמצעות קוד משתמש, המאפשר לקשר בין מנהל בית-הספר לתלמידים שנדגמו ובין כל אחד מהמורים שהשתתפו לבין תלמידיו. גם השאלון להורים היה מזהה באופן דומה באמצעות שם משתמש. שאלונים אלו נשלחו להורים בידי התלמידים שהשתתפו במחקר, והוחזרו במעטפה סגורה תוך שמירה על אנונימיות של ההורים כלפי בית-הספר.

בפירלס, המבחנים המודפסים והשאלונים לתלמיד הועברו בבתי-הספר בידי בוחנים חיצוניים מיומנים שאינם מצוות בית-הספר. הבוחנים גויסו, הועסקו והוכשרו לתפקידם על ידי מרכז המחקר בישראל, בהתאם להנחיות שסיפקו מארגני המחקר הבין-לאומי ועל פי המדריך לבוחן שנכתב במרכז המחקר הבין-לאומי ותורגם בישראל לעברית ולערבית. הבחינה הועברה בשעות הבוקר. בכל כיתה נכחו בעת המבחן שני בוחנים ומורה משגיח מתוך צוות בית-הספר. בתחילת הבחינה הקריאו הבוחנים לתלמידים הוראות אחידות (ההוראות, שתורגמו לשפות המבחן השונות, סופקו על ידי מארגני המחקר במדריך לבוחן). כל אחד מהתלמידים קיבל לידייו את אחת מגרסאות המבחן, שהוקצתה לו מראש באקראי באמצעות תוכנה של מרכז המחקר הבין-לאומי (ראה פירוט על חוברות המבחן בסעיף 3.1.1). לתלמידים הוקצה זמן כולל של שעה וחצי לקרוא את הטקסטים ולהשיב על שאלות המבחן - 40 דקות לכל מקבץ. אחר כך ניתנה להם הפסקה קצרה, ולאחריה הם נתבקשו למלא את השאלון לתלמיד. לשם כך הוקצו להם 30 דקות (ובמידת הצורך ניתנה הארכה של 10 דקות נוספות, כאשר לדוגמה היו תלמידים שלא סיימו למלא את השאלון).

באי-פירלס, המבחנים הממוחשבים והשאלון הקצר הועברו בבתי-הספר במועד שנקבע מראש עבור כל בית-ספר משתתף, בדרך כלל כשבוע לאחר העברת המבחן המודפס פירלס. צוות הבוחנים הגיע לבית-הספר כשעה לפני תחילת המבחן מצויד בכ-20 מחשבים ניידים, אשר הובאו ממרכז המחקר בישראל, עליהם הוותקנה מראש תוכנת המבחן. התלמידים בכיתת האם חולקו מראש לשתי קבוצות שנבחנו בזו אחר זו, וזאת בכדי להקטין את העומס בכיתת המחשבים ולאפשר הבחנות בתנאים אופטימליים, כאשר כל תלמיד ישוב ליד שולחן נפרד. שני הבוחנים, אשר האחד מהם היה מפעיל מחשבים והתמצא במערכות ממוחשבות והאחר בוחן מיומן, גויסו והוכשרו לתפקידם על ידי מרכז המחקר בישראל, בהתאם להנחיות שסיפק הארגון הבינלאומי. המחשבים הוותקנו בכיתת המבחן לפני כניסת התלמידים. בכיתה נכח גם מורה מבית-הספר או נציג אחר של בית-הספר שתפקידו היה לשמור על סדר ועל "אווירת בחינה" בכיתה. המורה לא התערב בהעברת המבחן ולא נחשף לחומרי המבחן. עם תחילת המבחן, כל תלמיד הושב מול מחשב נייד וזוהה על ידי הזנת קוד אישי ספציפי שהוכן מבעוד מועד. בתחילת המבחן הקריאו הבוחנים לתלמידים הוראות אחידות. כל אחד מהתלמידים קיבל אחת מגרסאות המבחן שהוקצתה לו מראש באופן אקראי לאחר הקראת הברכות וההוראות החל שלב היכרות עם סביבת המבחן הממוחשב שנמשכה כ-10 דקות ולאחריה החל המבחן עצמו שנמשך ארבעים דקות (20 דקות לכל טקסט). לאחר סיום המבחן התלמידים מלאו את

השאלון הקצר. בסוף כל יום נאספו קבצי התגובות של הנבחרים בכל כיתה והועברו למרכז המחקר, שהעביר אותם למרכז המחקר הבין-לאומי.

3.3.3 בקרת איכות על העברת המבחנים והשאלונים

נציגי ראמ"ה ערכו "ביקורי פתע" בעת המבחן בכמחצית מבתי-הספר המשתתפים, וזאת ללא תאום מראש עם בית-הספר. ככלל, המבחנים התקיימו כסדרם ועל פי נוהלי מרכז המחקר הבין-לאומי, ולא נצפו בעיות מיוחדות ו/או חריגות מן הנהלים. נוסף על כך מינה מרכז המחקר הבין-לאומי בִּקְר איכות מטעמו בכל מדינה משתתפת. הַבְּקָר הוכשר באמצעות חומרי קריאה ששלחו אליו מארגני המחקר הבין-לאומי ונכח ב-10% מבתי-ספר שנדגמו באקראי על ידי מרכז המחקר הבין-לאומי. הוא מילא דוחות מפורטים על התנהלות הבחינה בכל אחד מהם והעביר אותם ישירות למארגני מחקר פירלס. ככל הידוע, גם על פי דוחות אלו לא נרשמו בעיות חריגות בהעברת המבחנים בישראל.

3.3.4 בדיקת המבחנים וקידודם (ציינון)

המבחנים נבדקו בידי סטודנטים בעלי תואר ראשון לפחות בתחומים שונים של מדעי החברה, שמוינו, גויסו והוכשרו על ידי מרכז המחקר בישראל (ראמ"ה) לאחר שנמצאו מתאימים למשימה. ההכשרה כללה תרגול של בדיקת פריטי המבחן ודיון בתשובות שונות של תלמידים בכל הפריטים הפתוחים (דהיינו, בשאלות הפתוחות שבהן הנבחן נדרש להפיץ ולכתוב תשובה בכוחות עצמו ולא לבחור מתוך כמה תשובות נתונות). ההכשרה כללה גם למידה והבנה של המחונן (מדריך הקידוד), תרגול של תהליך הציינון מתוך דוגמאות של תשובות נכונות ושגויות של תלמידים, ולצדן הציון הנכון על פי המחונן וההגדרות של מרכז המחקר הבין-לאומי. הפריטים הפתוחים במבחנים נבדקו בהתאם להנחיות המפורטות במחונן של המבחנים, שנכתב באנגלית במרכז המחקר הבין-לאומי ותורגם במרכז המחקר בישראל לשתי שפות המטרה – עברית וערבית, בהליך תרגום מקצועי. מעריכי המבחנים לוו לאורך כל הבדיקה במעריכים בכירים, בעלי ניסיון עשיר בהוראת שפת אם בבית-הספר היסודי בכל אחת משפות המבחן. בדיקת המבחנים הייתה רוחבית (כלומר לפי טקסטים וכל הפרטים המקושרים אליהם ולא לפי חוברות המבחן: בעבור כל טקסט נערכה תחילה הדרכה של הבדיקה של הפריטים המקושרים אליו ולאחר מכן התבצעה בדיקה רוחבית של פריט אחרי פריט בכל חוברות המבחן בהן הוא הופיע). כמו כן, כרבע מתשובות התלמידים נבדקו בידי שני מעריכים בלתי תלויים לצורך בדיקת מהימנות בין מעריכים. הציונים (קודים) שניתנו על הפריטים הפתוחים והתשובות שהנבחנים סימנו בשאלות הסגורות במבחנים ובשאלונים הוקלדו לתוכנה ייעודית של מרכז המחקר הבין-לאומי. מאגר הנתונים של תשובות התלמידים בשאלות הסגורות והציונים שנתנו להם הבודקים על כל פריט הועברו עם תום הבדיקה למרכז הבין-לאומי.

במבחן אי-פירלס, הבדיקה נעשתה באתר אינטרנט ייעודי שפותח לצורך המשימה. כאמור, כל קבצי הנתונים של המבחן הועלו לאתר ייעודי של הארגון הבינלאומי. הארגון חילץ מקבצים אלו את תשובות התלמידים ו"שתל" אותן באתר הבדיקה. בסיומה של כל הדרכה, הוקצו לבודקים הפריטים שלגביהם התבצעה ההדרכה והם צפו וקודדו את תשובות התלמידים בזו אחר זו, בבדיקה רוחבית, כמתואר לעיל. גם בבדיקה זו שולב הליך של בדיקת מהימנות, כאשר רבע מתשובות התלמידים נבדקו במקביל על ידי שני בודקים.

3.3.5 טיוב הנתונים

טיוב הנתונים ועיבודם במחקר נעשו במרכז המחקר הבין-לאומי בשיתוף פעולה עם מנהלי המחקר במדינות המשתתפות. בדומה למחקר החלוץ, גם במחקר העיקרי הופק לכל מדינה דוח נתונים פסיכומטריים לכל פריטי המבחן והשאלונים, מתוך הנתונים שנאספו בה. בשלב זה התבקשו מנהלי המחקר בכל מדינה לבחון אם יש ממצאים מספריים חריגים שעשויים להעיד כי בשאלון או במבחן יש פריטים שלא תורגמו, הופקו או נבדקו כראוי ויש להוציא את נתוניהם מניתוח התוצאות הכולל.

בישראל אותרו מרבית הבעיות בתרגום ובהתאמה התרבותית עוד במחקר החלוץ, ותוקנו טרם המחקר העיקרי. לא נמצאו בעיות נוספות במחקר העיקרי.

3.4 עיבוד הנתונים

3.4.1 ניתוח הנתונים במחקר פירלס וחישוב הציונים

3.4.1.1 תאוריית התגובה לפריט

תיאוריית התגובה לפריט (IRT - Item Response Theory) היא הגישה התיאורטית המשמשת לניתוח נתוני המבחנים במחקר פירלס. בבסיס התיאוריה הזאת, בעבור כל פריט מוגדר טיב הקשר בין "רמת היכולת של הנבחן" (הידע שלו, שליטתו בתחום הדעת) ובין ההסתברות שיצליח להשיב נכונה על הפריט הנתון. אחת מהנחות היסוד המרכזיות בתיאוריית התגובה לפריט היא שכל הפריטים מודדים את אותה תכונה (במחקר זה - ידע או מיומנות בתחום הדעת הנבדק). הנחה נוספת היא שההסתברות להשיב נכונה על פריט מסוים אינה תלויה ביכולת להשיב נכונה על פריט אחר במבחן. אחד היתרונות של תיאוריה זו הוא שאפשר לחשב באמצעותה אומדני רמת יכולת לכל הנבחנים על אותו סולם, גם אם נבחנו שונים השיבו על נוסחים שונים של המבחן (כלומר השיבו על מקבצים שונים של פריטים), ובלבד שהפריטים השונים מודדים את אותה יכולת (או, במחקר זה, ידע באותו תחום דעת).

יתרון נוסף של שימוש בתיאוריית התגובה לפריט הוא שרמת היכולת של הנבחן ורמת הקושי של הפריטים מחושבות במונחי אותו סולם – הסולם של התכונה הנמדדת (במחקר זה – סולם רמת הידע והמיומנות בתחום דעת נתון). לפי מודל ה-IRT, הקשר בין הצלחתו של נבחן ובין רמת קושי הפריטים בא לידי ביטוי בכך שככל שהנבחן הוא בעל יכולת (ידע) גבוהה מרמת הקושי של הפריט הנתון, ההסתברות כי ישיב נכונה על הפריט גבוהה יותר; ככל שהנבחן הוא בעל יכולת (ידע) נמוכה מרמת הקושי של פריט נתון, ההסתברות שיוכל להשיב נכונה על הפריט נמוכה יותר.

במחקר פירלס משתמשים במודל ה-IRT החד-פרמטרי (מודל Rasch), בשילוב מודלים המותאמים לעיבוד נתוני פריטים פוליטומיים (Polytomous Items) – פריטים שיש בהם יותר משתי רמות תשובה, כגון פריטים שבהם התגובה יכולה להיות נכונה, נכונה חלקית או שגויה, והניקוד עליהם תלוי באיכות התשובה). במודל זה אין התייחסות לרמת ההבחנתיות של הפריט, ואף לא לרמת ההסתברות לנחש את התשובה הנכונה לגביו, גם אם לנבחן אין כל ידע בתחום הדעת הנבדק. יתרון של מודל זה הוא בפשטותו יחסית למודלים מורכבים יותר של IRT.

לאחר איסוף הנתונים, מרכז המחקר הבין-לאומי אומד תחילה את מאפייני הפרמטרים של הפריטים בהליך המכונה "כיול הפרמטרים של הפריטים": בעבור כל פריט נאמדת רמת הקושי של כל אחת מרמות התשובה האפשריות. פעולה זו נעשית בנפרד בסוגות הטקסט השונות ובמיומנויות הקריאה השונות. על בסיס אומדנים אלו של מאפייני הפריטים ועל פי תגובות הנבחנים על פריטי המבחן השונים, אומדים בהמשך את רמת הידע של כל נבחן ונבחן. בהליך כיול הפרמטרים משולבים גם תגובותיהם של התלמידים במחזור מחקר זה ובמחזורים קודמים ל"פריטי העוגן" (פריטים מקשרים, המשותפים למחקר פירלס הנוכחי ולמחזורי המחקר הקודמים), מה שמאפשר לכייל את הפרמטרים של הפריטים במחקר 2016 לסולם מחקר העבר, ואגב כך גם לכייל את ציוני התלמידים במחזור זה לסולם הציונים הכללי במחזורים קודמים של מחקר פירלס. כיול זה מאפשר השוואה של הציונים לאורך מחזורי מחקר שונים.

מבחן אי-פירלס נערך לראשונה ב-2016 ב-14 המדינות שהשתתפו בשני המבחנים, פירלס ואי-פירלס. צימוד הנתונים משני המבחנים וקיבוע הפרמטרים במבחן פירלס אפשרו לאמוד את הפרמטרים של הפריטים במבחן אי-פירלס על סולם הציונים הכללי של מבחן פירלס. שימוש במודלים IRT רב-ממדיים אפשר להפיק ציונים במבחן הממוחשב אי-פירלס, והללו מוקמו על גבי סולם הציונים

הכללי של המבחן המודפס פירלס באמצעות נוסחאות כיוול. כך מתאפשרת השוואה בין ההישגים בשני המבחנים במחקר, בקריאה של טקסטים מודפסים לעומת קריאה בסביבה ממוחשבת.

3.4.1.2 ערכים סבירים (plausible values)

מחקר פירלס נועד לאמוד את התפלגות רמת היכולת (הידע) של כלל אוכלוסיית המטרה, קרי התלמידים בדרגת הכיתה הרלוונטית בכל מדינה שמשתתפת במחקר, או של קבוצת נבחנים רחבה ולא את רמת היכולת של נבחן יחיד. כאמור, מחד אין פריטי מבחן הניתנים לכל הנבחנים ומאידך כל נבחן מקבל רק מספר קטן יחסית של פריטי מבחן מבין כלל פרטי המבחן. בשל כך, אומדן רמת היכולת של הנבחן היחיד נתון להיקף מסוים של טעות או אי-ודאות המדידה. כיוון שכך, אומדני הפרמטרים (ממוצע, סטיית תקן ואחוזונים) של התפלגות רמת היכולת של כלל האוכלוסייה, המתבססים ישירות על אומדני רמת היכולת של הנבחנים היחידים, עלולים להיות מדויקים פחות. זאת ועוד, אומדני הפרמטרים של רמת קושי הפריטים ואומדני היכולת של כל נבחן מתבססים על אותו מאגר של תגובות הנבחנים לפריטי המבחן, ועל כן קיימת התאמה גבוהה (התאמת יתר) ביניהם, המתבטאת בין היתר בסטיית התקן/השונות של הציונים קטנה יותר. כדי להתגבר על קשיים אלו, משתמשים במחקרים רחבי היקף (כדוגמת מחקרי פירלס, טימס, פיזה, טאליס ואף סקר המיומנויות של האוכלוסייה הבוגרת PIAAC) במתודולוגיה שבה אומדים את הפרמטרים של התפלגות רמת היכולת של כלל אוכלוסיית התלמידים הרלוונטית תוך שימוש בערכים מרובים שמייצגים התפלגות אפשרית של רמת היכולת של כל נבחן. לצורך אמידת הממוצע והשונות של רמת היכולת של האוכלוסייה, ולצורך אמידת פרמטרים אחרים של ההתפלגות (כגון אחוזונים), דוגמים באקראי בעבור כל נבחן חמישה ערכים מתוך פונקציית ההתפלגות הנאמדת המותנית של רמת היכולת שלו. חמישה ערכים אלו, המייצגים חמישה אומדנים של רמת היכולת של נבחן נתון, מכונים "ערכים סבירים" (plausible values, או בקיצור PV), והם משמשים לחישוב המדדים המסכמים של כלל אוכלוסיית הנבחנים. לשם שיפור אומדני הפרמטרים ואומדני הציונים של הנבחנים משתמשים בהליך האמידה של הפרמטרים לא רק בתגובות הנבחנים על כל אחד מהפריטים שנכללו במבחן שקבלו (כלומר הקודים או הציונים שנתנו לנבחן בהתאם לביצוע שלו בכל פריט ופריט במבחן), אלא גם בכל מידע אחר שנאסף על אודות הנבחנים ואשר עשוי לסייע באמידת ההתפלגות של רמת יכולתם, כגון נתוני רקע השאובים מהשאלונים שמילאו. חשוב להדגיש שאין להתייחס אל הערכים הסבירים כאל אינדיקטורים של ציון או אומדן יכולת של הנבחן היחיד.

3.4.1.3 סולם הציונים

כאמור, עקב הליך כיוול הפרמטרים של הפריטים, תוך שימוש ב"פריטי עוגן", ושל רמת היכולת (רמת הידע) של כל נבחן ונבחן, הציונים המתקבלים במבחן מכילים לסולם הציונים של פירלס במחזורי קודמים. משמעות הדבר היא שהציונים הנוכחיים הם בני השוואה לציונים במחזורי מחקר קודמים. סולם הציונים המקורי של מחקר פירלס נקבע במחזור המחקר הראשון בשנת 2001 כסולם ציונים רציף כך שממוצע הציונים בכל המדינות שהשתתפו באותה שנה¹³ היה 500 וסטיית התקן הייתה 100. תיאורטית, הציונים יכולים לנוע בין 0 ל-1,000, אך בפועל, טווח הציונים הוא בין 200 ל-800 בעבור נבחן יחיד (ובעבור ממוצע המדינות הוא אף מצומצם הרבה יותר).

3.4.1.4 רמות הישג

כדי לתת משמעות נוספת לציונים, מרכז המחקר הבין-לאומי קבע כמה "רמות הישג". לאחר החלטה על מספר הרמות נקבעו סְפֵי ההישגים (Benchmarks) על סולם ציוני פירלס – סְפֵי ההישג הם ערכים על סולם הציונים שמסמנים את הגבול בין רמה לרמה. מדובר בהליך שיטתי המביא בחשבון את רמת הקושי של שאלות המבחן ואת המיומנות והידע הנדרשים מן התלמיד בכל אחת מרמות

¹³ משמעות השינוי בהרכב המדינות המשתתפות היא שממוצע ההישגים הבין-לאומי בעבור כל תחום תוכן משתנה מעט בכל פעם, מכיוון שהוא מחושב על פי הרכב המדינות המשתתפות באותו מחזור מחקר.

ההישג כדי להתמודד בהצלחה עם השאלות. במחקר פירלס 2016, בדומה למחזורי פירלס קודמים,¹⁴ נקבעו חמש קטגוריות של רמת ההישג:

(1) הישג ברמת "הצטיינות" (Advanced Benchmark) – בקטגוריה זו נמצאים תלמידים שסוף הישגיהם במבחן הוא 625 נקודות (כלומר 625 נקודות או יותר);

(2) הישג ברמה "גבוהה" (High Benchmark) – בקטגוריה זו נמצאים תלמידים שסוף הישגיהם במבחן הוא 550 נקודות (ועד 625 נקודות);

(3) הישג ברמה "בינונית" (Intermediate Benchmark) – בקטגוריה זו נמצאים תלמידים שסוף הישגיהם במבחן הוא 475 נקודות (ועד 550 נקודות);

(4) הישג ברמה "נמוכה" (Low Benchmark) – בקטגוריה זו נמצאים תלמידים שסוף הישגיהם במבחן הוא 400 נקודות (ועד 475 נקודות);

(5) הישג ברמה "מתחת לסף" – בקטגוריה זו נמצאים תלמידים שהישגיהם במבחן נמוכים מ-400 נקודות.

בהליך המכונה "עיון הסולם" (Scale Anchoring), פריטי המבחן מסווגים במרכז המחקר הבין-לאומי לרמות ההישג לפי רמת הקושי שלהם. בד בבד, כל נבחן מסווג לקטגוריה המייצגת את רמת ההישג הגבוהה ביותר שבה הוא מצופה לענות נכונה על מרבית השאלות המתאימות לאותה רמה. המיון מתבסס על ההנחה שלפיה תלמיד ברמת הישג נתונה צפוי להשיב נכונה על מחצית לפחות משאלות המבחן ברמה זו. מובן מאליו שלא מכל התלמידים ברמת הישג נתונה מצפים לשיעור זה של תשובות נכונות. תלמיד שהישגיו בגבול התחתון של אותה רמת הישג צפוי לענות נכונה על 50% מן השאלות ברמה זו, ולעומתו תלמיד שהישגיו קרובים לגבול העליון של רמת ההישג, צפוי לענות נכונה על שיעור גבוה יותר של שאלות ברמה זו. תלמיד שהישגיו בקטגוריה הגבוהה ביותר (ההישג ברמת "הצטיינות") צפוי לענות נכונה על רוב פריטי המבחן.¹⁵ בשלב הבא, מומחי תוכן מנתחים ונותנים פשר לתכנים ולדרישות הקוגניטיביות המאפיינים את הפריטים ששויכו לכל רמת הישג, ובעקבות זאת מוגדרים הידע והכישורים הנדרשים מהתלמידים שמסוגלים לענות נכונה על פריטים ברמת הישג נתונה.

תיאור תמציתי של מיומנויות הקריאה הנדרשות בכל אחת מרמות ההישג מוצג בלוח 3.5 בעבור קריאה של טקסטים מודפסים במבחן פירלס, ובלוח 3.6 בעבור קריאה בסביבה ממוחשבת במבחן אי-פירלס.

¹⁴ במחזור המחקר הראשון, בשנת 2001, ספיי ההישג שנקבעו היו הציונים הבין-לאומיים של האחוזונים ה-90, ה-75, ה-50 וה-25, שאפשרו להגדיר את העשירון העליון, הרבעון העליון, החציון העליון וטווח הציונים שמעל האחוזון העליון. החל ממחזור מחקר 2006 נקבעו 4 ערכי סף אחידים (כאמור: 625, 550, 475 ו-400) והם משמשים מאז ועד היום לקביעת רמות ההישג בקריאה. בהצגת מגמות לאורך מחזורי מחקר נעשה שימוש בספים אלו.

¹⁵ הרחבה על שיטת הפיתוח של רמות ההישג אפשר למצוא במדריך הטכני של דוח פירלס 2016, המפרט את השיטות והנהלים שיושמו במחקר: <https://timssandpirls.bc.edu/publications/pirls/2016-methods.html>

לוח 3.5: תיאור תמציתי של חמש רמות הישג במבחן פירלס

רמת הישג	סף הישג	מה התלמידים מסוגלים לעשות
"הצטיינות"	625 נקודות	תלמידים ברמת הישג זו מסוגלים לקרוא טקסט יחסית מורכב. בקוראם טקסט ספרותי, הם מסוגלים לשלב מידע הפזור בטקסט ולספק פרשנות מבוססת טקסט בנוגע לאופיו של הגיבור, כוונותיו ורגשותיו, וכיצד הללו משתנים לאורך הטקסט. הם מבינים את תפקידם של מארגני הטקסט וגם מתחילים להעריך את ההשפעה של הסגנון והשפה בהם בחר המחבר על הקורא. בקוראם טקסטים מידעיים הם יכולים לזהות ולפרש מידע מחלקים שונים של הטקסט, ולספק תמיכה לקביעותיהם מחלקים שונים בו. הם יכולים לשלב מידע מחלקיו השונים כדי לתאר רצף פעילות ולהסביר יחסים, והם גם מתחילים להעריך מרכיבים מילוליים וחזותיים ואף לשקול את עמדתו ונקודת מבטו של המחבר.
"גבוהה"	550 נקודות	תלמידים ברמת הישג זו מסוגלים לקרוא טקסט יחסית מורכב. הם מסוגלים להבחין לאזכר ולשלב פרטים משמעותיים ורעיונות הפזורים בטקסט, ויכולים לספק להסקותיהם תמיכה המעוגנת בטקסט. הם יכולים לפרש ולשלב בין פרטים ואירועים, ובכלל זה מאפייני דמויות ורגשות, ככל שאלו מתפתחים ומשתנים לאורך הטקסט. הם גם יכולים לזהות ולהשתמש במאפיינים טקסטואליים ומרכיבי טקסט (מטאפורה, נימה, דימיון). בנוסף, בטקסטים מידעיים הם מסוגלים להשתמש בתכונות מארגנות כדי לנווט בו את עצמם, להסיק מסקנות ולמצוא קשרים, לשלב בין מידע מילולי וחזותי, ואף להעריך ולהכליל.
"בינונית"	475 נקודות	תלמידים ברמת הישג זו מסוגלים לקרוא טקסט פשוט משולב בטקסט יחסית מורכב. הם מסוגלים לזהות למקם ולשחזר אירועים מרכזיים, רצפי אירועים, ופרטים רלוונטיים אחרים המצוינים מפורשות בטקסט, הם יכולים להסיק הסקות ישירות מהטקסט בנוגע לתכונות, רגשות ומוטיבציות של דמויות מרכזיות ומקשרים בין חלקים שונים בו. הם יכולים לפרש קשרים ברורים של סיבה ותוצאה, לזהות עדויות ולספק דוגמאות, ולעיתים אף להתחיל לזהות מאפיינים טקסטואליים ומרכיבי טקסט. בנוסף, בטקסטים המידעיים הם מסוגלים לאתר ולשחזר 2 עד 3 פיסות מידע מן הטקסט, ולספק הסבר עובדתי.
"נמוכה"	400 נקודות	תלמידים ברמת הישג זו מסוגלים לקרוא טקסט פשוט והם מגלים מיומנויות קריאה ואסטרטגיות ברמה בסיסית למדי. הם מסוגלים לזהות, למקם ולאחזר פרטים המצוינים מפורשות בטקסט, בעיקר אם אלה ממוקמים בראשית הטקסט, ולעיתים הם יכולים להסיק הסקות ישירות מהטקסט, ואף לפרש רעיונות מרכזיים.
"מתחת לסף"		תלמידים ברמת הישג זו הם מתחת לסף הניתן להערכה – אין תיאור של ביצועיהם.

לוח 3.6: תיאור תמציתי של חמש רמות הישג במבחן אי-פירלס (קריאת טקסט מידעי בלבד)

רמת הישג	סף הישג	מה התלמידים מסוגלים לעשות
"הצטיינות"	625 נקודות	תלמידים ברמת הישג זו מסוגלים להסיק הסקות ממידע מורכב בכדי להסביר ולתמוך בטענותם. הם יכולים לפרש ולשלב בין פרטים הפזורים בדף אתר ואף בין דפים שונים בכדי להסביר יחסים בין פרטים, תוך הצגת הבנה מעמיקה. הם גם יכולים להעריך את ההשפעה של מרכיבים מילוליים, חזותיים ואינטראקטיביים ואף שוקלים לעיתים את עמדתו ונקודת מבטו של המחבר.
"גבוהה"	550 נקודות	תלמידים ברמת הישג זו מסוגלים להבחין בפרטים משמעותיים ולערוך השוואות, הם יכולים לפרש ולשלב בין פרטים הפזורים בדף אתר ואף בין דפים שונים בכדי לספק דוגמאות ולהנגיד בין פרטים. הם גם יכולים לזהות ולהעריך כיצד מאפיינים טקסטואליים ומרכיבי טקסט כמו גם מרכיבים גרפיים יכולים לחזק ולתמוך בתוכן המוצג.
"בינונית"	475 נקודות	תלמידים ברמת הישג זו מסוגלים לזהות למקם ולשחזר מידע המוצג במגוון צורות, ובכלל זה שימוש עצמאי בכלי ניווט. הם מסוגלים להסיק הסקות בכדי לזהות סיבות ופעולות, לפרש ולשלב מידע הפזור בדף האתר לשם השוואה ומתן הסר או זיהוי קשר סיבתי. הם לעיתים גם מסוגלים להעריך את השימוש במאפיינים אינטראקטיביים בכדי לרכוש מידע.
"נמוכה"	400 נקודות	תלמידים ברמת הישג זו מסוגלים לזהות, למקם ולשחזר פרטים המצוינים מפורשות בדפי האתר, אשר מכילים טקסט, מגוון כלי ניווט פשוטים ומאפיינים דינמיים (חלונות קופצים, צירי זמן וכו'). לעתים הם יכולים להסיק הסקות ישירות לגבי תיאורים.
"מתחת לסף"		תלמידים ברמת הישג זו הם מתחת לסף הניתן להערכה – אין תיאור של ביצועיהם.

3.4.2 עמדות התלמידים כלפי קריאה

בירור עמדות התלמידים כלפי קריאה נעשה באמצעות שאלון שהועבר להם מיד עם סיום המבחן. כאמור, למילוי השאלון הוקצו כ-30 דקות (ובמידת הצורך ניתנה הארכה של 10 דקות נוספות, כאשר לדוגמה היו תלמידים שלא סיימו למלא את השאלון). השאלון כלל, בין היתר, מספר שאלות לגבי כל אחד משלושה היבטים: אהבה לקרוא, הביטחון ביכולת לקרוא, הוראה מעודדת מעורבות בשיעורי שפה (לפירוט ראו פרק 6.1). השאלות נוסחו כהיגדים והתלמיד התבקש לציין באיזו מידה הוא מסכים עם כל היגד על סולם של ארבע דרגות (1: מאוד לא מסכים, 2: קצת לא מסכים, 3: קצת מסכים, ו-4: מאוד מסכים), או לדווח על תדירות הקריאה מחוץ לכותלי בית הספר (1: בכלל לא או כמעט בכלל לא, 2: פעם או פעמיים בחודש, 3: פעם או פעמיים בשבוע, ו-4: בכל יום או כמעט בכל יום). תגובותיו של תלמיד לכלל ההיגדים בהיבט נתון, או מידת הסכמתו עימם, שוקללו במדד אישי. כך, עבור כל תלמיד, ככל שמידת הסכמתו עם קבוצת היגדים בהיבט נתון גבוהה יותר כך ערך המדד האישי גבוה יותר – ועל כן הדבר משקף עמדה חיובית יותר של התלמיד בהיבט הנבדק. על "המרת" תגובות התלמידים להיגדים השונים בכל היבט לכדי מדד מסכם וחלוקתו לשלוש דרגות (עמדות חיוביות מאוד, עמדות די חיוביות, ועמדות שליליות) ראו הרחבה בפרק 6.1.

בכל מדינה חושב ממוצע מדד עבור כל היבט, בהתבסס על המדדים האישיים של כלל התלמידים בהיבט הנבדק. להרחבה על קביעת סולמות המדדים בכל עמדות התלמידים כלפי קריאה ראו תיבה

3.4

תיבה 3.4: קביעת סולמות המדדים בכל עמדות התלמידים כלפי קריאה

בהתבסס על התפלגות תגובות התלמידים עם ההיגדים בכל היבט נבדק בקרב המדינות המשתתפות במחקר פירלס 2016, ובמטרה לספק נקודת ייחוס להשוואה בין המדינות, המדדים עברו טרנספורמציה כך שבכל מדד נקבע ממוצע המדינות המשתתפות ל-10 נקודות (בסולם 0 עד 20), וסטיית התקן על 2 נקודות. על הסולם קבעו מארגני המחקר שלוש דרגות לכל מדד, המשקפות את מידת ההסכמה עם ההיגדים המרכיבים את המדד: וסימן "מאוד" (עמדה חיובית), "די" (עמדה די חיובית), ו-"לא" (עמדת שלילית) בסמיכות לשם המדד (כך עבור המדד "אוהב" הדרגות הן: מאוד אוהב, די אוהב ולא אוהב, ובהתאם לכך עבור שאר המדדים). השיקולים לקביעת ערכי הסף לחלוקת הסולם הרציף של ציוני המדד לדרגות היו שונים בכל אחד משלושת המדדים (בהתאם, למשל, למספר ההיגדים המרכיבים את המדד). **בדומה לאופן נקבעו מדדים הנוגעים לעמדות תלמידים, נקבעו גם המדדים הנוגעים לדיווחי מורים וכן אלו המושגים על דיווחי ההורים.** לפירוט כיצד נקבעה החלוקה בכל אחד מהמדדים ראו פרק השיטות והפרוצדורות באתר פירלס הבין-לאומי¹⁶ הנוגע לכלל המדדים בשאלוני ההקשר (של תלמידים, של הוריהם, של מורים ושל מנהלים) וכן בתיבות בפרקים 6 עד 8 להלן.

בין מחזור מחקר קודם (2011) לנוכחי (2016), שאלוני המחקר, ובהם גם השאלון לתלמיד, עברו שינויים ועדכונים וזאת במטרה לשפר את המדידה, ובפרט את התוקף והמהימנות של המדדים. כך, למדדים מסוימים נוספו היגדים המרכיבים אותם, בכדי לבחון ולהעמיק בסוגיות רלוונטיות נוספות, או שהנוסח של היגדים מסוימים שופר. ככלל, השוואת מדדים בין מחזורי המחקר אפשרית רק במידה והתקיימו מספר תנאים: א. לפחות 5 היגדים זהים משותפים לשני מחזורי המחקר; ב. ההיגדים המשותפים מהווים למעלה ממחצית ההיגדים לחישוב המדד במחזור 2016; ו-ג. מאפייני מדידה דומים של המדד וההיגדים המשותפים בשני מחזורי המחקר. תנאים אלו לא התקיימו בעבור מדדי עמדות התלמידים, ועל כן הן באתר המחקר והן בדוח הישראלי לא מוצגות בעבורם מגמות לאורך שנים.

3.4.3 עמדות ותפיסות המורים לשפה

המידע בפרק 7 מבוסס כאמור על התשובות בשאלון שהועבר למורי הקריאה (שפה) בכיתות שנדגמו למחקר. מכל בית-ספר נדגם למעשה מורה אחד, וסה"כ 159 מורים ענו על השאלונים (כמפורט בלוח 3.4). חשוב להדגיש כי בשל אופייה של שיטה זו של דגימת המורים למחקר, אין לראות במורים שמילאו את השאלונים מדגם מייצג של כלל המורים לקריאה בכיתות ד' בישראל (כך גם בשאר המדינות המשתתפות במחקר). על אף הסתייגות זו, הנתונים מאפשרים לבחון את מאפייני המורים המלמדים שפה בכיתות בהם נדגמו תלמידים להשתתף במחקר בישראל (ואשר הישגי תלמידים אלו בקריאה נמדדו במחקר). לנוכח העובדה שהמורים שמילאו את השאלונים אינם בהכרח מדגם מייצג,

¹⁶ פרק השיטות והפרוצדורות באתר מחקר פירלס, בקישור:

https://timssandpirls.bc.edu/publications/pirls/2016-methods/P16_MP_Chap14_Context_Questionnaire_Scales.pdf

לפירוט המדדים, ערכי הסף בין הקטגוריות השונות בכל מדד, ראו בפרק הממצאים באתר מחקר פירלס, לדוגמא בקישורים הבאים:

שאלוני תלמידים הנוגעים לעמדות תלמידים:

<http://timssandpirls.bc.edu/pirls2016/international-results/pirls/student-engagement-and-attitudes/>

שאלוני מורים הנוגעים למאפייני הרקע של המורים, שביעות רצון, אקלים בית ספרי וסביבה ופדגוגית:

<http://timssandpirls.bc.edu/pirls2016/international-results/pirls/teachers-and-principals-preparation/>

<http://timssandpirls.bc.edu/pirls2016/international-results/pirls/school-climate/>

<http://timssandpirls.bc.edu/pirls2016/international-results/pirls/school-safety/>

שאלוני הורים הנוגעים לסביבה הביתית, חינוך לגיל הרך ועמדות הורים:

<http://timssandpirls.bc.edu/pirls2016/international-results/pirls/home-environment-support/>

חשוב לציין גם שהנתונים בפרק זה מנותחים ברמת תלמידים ולא ברמת מורים, כלומר, מוצג שיעור התלמידים שלמדו אצל מורים בעלי מאפיין נתון, ולא שיעור המורים בעלי מאפיין זה (למשל, שיעור התלמידים שלמדו אצל מורים שהשתתפו בפעילות לפיתוח מקצועי בשנה שבה נערך המחקר).

השאלון למורים הועבר בפורמט ממוחשב באמצעות רשת האינטרנט. המורים התבקשו למלאו עוד לפני העברת המבחנים לתלמידים. מילוי השאלון ארך כ-30-45 דקות לכל מורה. כדי שיהיה אפשר לקשר בין המידע על אודות המורים לבין הישגי תלמידיהם, נאסף מידע המקשר בין כל אחד מהמורים לבין התלמידים שהוא מלמד.

בשאלות העוסקות בעמדות המורים ובמאפייניהם התבקש המורה לציין באיזו מידה הוא מסכים עם כל היגד נתון על סולם של 4 דרגות (1: מאוד לא מסכים, 2: קצת לא מסכים, 3: קצת מסכים, ו-4: מאוד מסכים; או סולם דומה). שאלות העוסקות בעמדות הורכבו על פי רוב מקבוצת היגדים רלוונטית לנושא הנבדק באותה שאלה. תגובותיו של המורה לכלל ההיגדים בהיבט נתון (ולעתים תשובותיהם בכמה היגדים שהשתייכו לשאלות בנושאים אחרים), או מידת הסכמתו עימם, שוקללו במדד אישי. ערכי המדד היו על סולם רציף שעליו קבעו מארגני מחקר פירלס על פי רוב שלוש דרגות. עבור כל מורה, ככל שמידת הסכמתו עם קבוצת היגדים בהיבט נתון גבוהה יותר כך ערך המדד האישי גבוה יותר – ועל כן הדבר משקף עמדה חיובית יותר של המורה בהיבט הנבדק (למשל, תפיסה חיובית של המשיב בנוגע לנושא המוצג). בכל מדינה חושב ממוצע המדד עבור כל היבט, בהתבסס על המדדים האישיים של כלל המורים בהיבט הנבדק. קביעת סולמות המדדים עבור עמדות מורים דומה לזו עבור עמדות תלמידים, כמתואר בתיבה 3.4. בהתבסס על התפלגות תגובות המורים עם ההיגדים בכל היבט נבדק בקרב המדינות המשתתפות במחקר פירלס 2016, ובמטרה לספק נקודת ייחוס להשוואה בין המדינות, המדדים עברו טרנספורמציה כך שבכל מדד נקבע ממוצע המדינות המשתתפות ל-10 נקודות (בסולם 0 עד 20), וסטיית התקן על 2 נקודות. על "המרת" תגובות המורים להיגדים השונים בכל היבט לכדי מדד מסכם וחלוקתו לדרגות ראו פירוט בפרק 7.2. לצד אלו נקבעו גם מדדים פשוטים יותר, המבוססים על קיבוץ קטגוריות של תשובות מורים, כגון המדדים: 'השכלה פורמלית', 'ותק בהוראה' ו-'פיתוח מקצועי', כמתואר בפרק 7.1.

3.4.4 עמדות ותפיסות הורי התלמידים

המידע בפרק 8 מבוסס כאמור על התשובות בשאלון שהועבר להורי התלמידים בכיתות שנדגמו למחקר. מכל תלמיד התקבל שאלון אחד שמולא על ידי הוריו (או אדם אחר שאחראי לו). כאמור בסעיף 3.1.5, הוריהם של כ-78% מהתלמידים שהשתתפו במחקר השיבו על השאלון. הנתונים מאפשרים לבחון את המאפיינים של הורי התלמידים שנדגמו להשתתף במחקר (ואשר הישגי תלמידים אלו בקריאה נמדדו במחקר), אך יש לציין כי הנתונים מנותחים ברמת התלמידים ולא ברמת ההורים. כלומר, מוצג שיעור התלמידים שלהוריהם מאפיין נתון או עמדה מסוימת, ולא שיעור ההורים בעלי מאפיין זה (למשל שיעור התלמידים שהוריהם בעלי השכלה אקדמית).

ביום בו התלמידים נבחנו במבחן ההישגים בקריאה של טקסטים מודפסים, ההורים קבלו מעטפה ובה שאלון, והם התבקשו למלא את השאלון ולהחזירו למעטפה. המעטפה הוחזרה לעורכי המחקר ביום בו התלמידים נבחנו במבחן ההישגים בקריאה של טקסטים בסביבה ממוחשבת, אשר בדרך כלל נערך כשבוע לאחר מכן. משך הזמן המשוער המוקדש למילוי השאלון הוא 15 עד 30 דקות. כדי שיהיה אפשר לקשר בין המידע על אודות ההורים לבין הישגי ילדיהם, נאסף מידע המקשר בין כל הורה שמילא את השאלון לבין הילד שלו.

בירור עמדות ההורים ומאפייניהם נעשה באמצעות שאלון שכלל מספר שאלות לגבי מגוון היבטים, וביניהם: הסביבה הביתית, חינוך בגיל הרך, ועמדות ההורים כלפי קריאה והשכלה (לפירוט ראו פרק 8). השאלות נוסחו כהיגדים, וההורה התבקש לציין באיזה מידה הוא מסכים עם כל היגד על סולם של ארבע דרגות (1: מאוד לא מסכים, 2: קצת לא מסכים, 3: קצת מסכים, ו-4: מאוד מסכים; או סולם דומה). שאלות העוסקות בעמדות הורכבו על פי רוב מקבוצת היגדים רלוונטית לנושא הנבדק באותה

שאלה. תגובותיו של ההורה לכלל היגדים להיבט נתון (ולעיתים תשובותיהם בכמה היגדים שהשתייכו לשאלות בנושאים אחרים), או מידת הסכמתו עימם, שוקללו במדד אישי. ערכי המדד היו על סולם רציף שעליו קבעו מארגני מחקר פירלס על פי רוב שלוש דרגות. עבור כל הורה, ככל שמידת הסכמתו עם קבוצת היגדים בהיבט נתון גבוהה יותר כך ערך המדד האישי גבוה יותר – ועל כן משקף עמדה חיובית יותר של ההורה בהיבט הנבדק. בכל מדינה חושב ממוצע המדד עבור כל היבט, בהתבסס על המדדים האישיים של כלל ההורים בהיבט הנבדק. קביעת סולמות המדדים עבור תגובות ההורים וכן הטרנספורמציה שהמדדים עברו דומות לאלו שתוארו לעיל עבור מורים ועבור מדדי עמדות התלמידים. על "המרת" תגובות ההורים להיגדים השונים בכל היבט לכדי מדד מסכם וחלוקתו לדרגות ראו פירוט בפרק 8. לצד אלו נקבעו גם מדדים פשוטים יותר, המבוססים על קיבוץ קטגוריות של תשובות ההורים, כגון המדדים 'ביקור בגן הילדים' ו-'ציפיות ההורים להשכלת ילדם'.

3.5 דיווח התוצאות

התוצאות הבין-לאומיות של מחקר פירלס מתפרסמות באמצעי התקשורת, באינטרנט ובדוחות הרשמיים של מארגני המחקר הבין-לאומי. התוצאות של ישראל וניתוחים ייחודיים לישראל מתפרסמים בדוח זה ובאתר ראמ"ה. בדוח זה מוצגים הישגי התלמידים בישראל (ממוצע הישגים ורמות הישג) במבט בין-לאומי ובמבט ישראלי. במבט בין-לאומי, הישגי ישראל מוצגים בהשוואה ל-49 מדינות נוספות שהשתתפו במבחן פירלס, ול-13 מדינות נוספות שהשתתפו במבחן אי-פירלס. במבט פנים-ישראלי מוצגים הישגיהם של כלל התלמידים, וכן בפילוח לפי מגזר שפה, מגדר, רקע חברתי-כלכלי אישי וסוג פיקוח, ולפי שילוב של משתנים אלו, וזאת בכל מבחן בנפרד. נוסף על כך, ההישגים מדווחים לפי סוגות טקסט (במבחן פירלס בלבד) ומיומנויות הקריאה הנדרשות (בשני המבחנים) על פי המוגדר במסגרות המושגיות (ראו הרחבה עליהן בפרק 2). מגמות שינוי בהישגים לאורך מחזורי המחקר השונים מוצגות אף הן.

לשם העמקה בנתונים וביצוע מחקר המשך המבוססים על נתוני מחקר פירלס בישראל, הנתונים הגולמיים (כולל מידע רקע נוסף) מונגשים לחוקרים באקדמיה המעוניינים בכך (לאחר פנייה למרכז המחקר הלאומי). מסד הנתונים הבין-לאומי (ובכללם השאלונים והמבחנים) נגיש באתר פירלס לכל המעוניין.¹⁷

¹⁷ ראה פירוט בקישור: <https://timssandpirls.bc.edu/pirls2016/international-database/index.html>