

מדידה בשירות הלמידה: הלכה ומעשה

פרופ' מיכל בלר
מנכ"לית ראמ"ה

מרס 2013

ראמ"ה: מדידה בשירות הלמידה

ראמ"ה, בית אבגד, ז'בוטינסקי 5, רמת גן, 5252006 • טל': 03-5205503, פקס: 03-5205509

דוא"ל: rama@education.gov.il • אתר: <http://rama.education.gov.il>

תוכן העניינים

3	הקדמה
3	מבחנים רחבי היקף במערכות חינוכיות ותדירותם
5	עדכון מתכונת המדידה וההערכה: שילוב של הערכה חיצונית ופנימית
6	הערכה חיצונית
6	המיצ"ב
8	מה אפשר ללמוד מהמיצ"ב
8	מגמות בהישגים לאורך זמן והשוואת ציונים
10	הערכה ארצית מדגמית
10	המחקרים הבינלאומיים
11	משוב ארצי מדגמי
11	ניטור ארצי מדגמי של רמת האלימות בבתי הספר
12	הערכה בית-ספרית פנימית
13	מיצ"ב פנימי
13	מבחני מדף, מבחנים מעצבים ומאגרי משימות
16	רכז הערכה בית-ספרי
17	הערכת סגלי הוראה וחקירת פרקטיקות הוראה
17	הערכת סגלי הוראה
18	סקר מורים בינלאומי
18	על השימוש בנתוני מבחנים וסקרים לצורכי מחקר והערכת תכניות ופרויקטים
21	מדידה בשירות הלמידה - סיכום
23	ביבליוגרפיה

מדידה בשירות הלמידה

פרופ' מיכל בלר, מנכ"לית ראמ"ה

הקדמה

עיון במסמכי היעדים של שרי חינוך לדורותיהם מעלה כי נכללים ביניהם שלושה יעדי-על למערכת החינוך: מיצוי הפוטנציאל (הלימודי, היצירתי והערכי) של כל תלמיד, צמצום פערים וקיום סביבת לימוד בטוחה.

בהנחה שיש הסכמה רחבה על יעדים אלו, עולה השאלה כיצד נדע אם הושגו. כיצד יידע כל הורה אם אכן סיפקה מערכת החינוך לילדיו את הכלים הנחוצים לו כדי להשתלב בהצלחה כאזרח בוגר ופעיל בחברה? כיצד יידע כל אחד מהשותפים לתהליך החינוכי (מורים, מנהלי בתי ספר, ובעלי התפקידים השונים במערכת החינוך) אם מילא את תפקידו כהלכה ואם ניתן מענה הולם לתלמידים מסוגי רקע שונים? כיצד נאתר פערים וכיצד נדע אם הצטמצמו? כיצד יידע הציבור שדור העתיד של ילדי ישראל הוכשר לעמוד באתגרי המאה ה-21? איך יובטח לציבור כי המשאבים הרבים המופנים למערכת החינוך מנוצלים נכונה ומשפיעים כפי שתוכנן? ואיך יוכח למקבלי ההחלטות שכדאי למדינה להגדיל את השקעותיה בחינוך, אפילו על חשבון צרכים חשובים אחרים? כדי לבחון את מידת השגתם של היעדים הללו נדרשים כלי מדידה והערכה תקפים, הנבנים באופן מקצועי.

מדידה והערכה הן סוגיות מורכבות בכל המגזרים - במגזר העסקי ואף יותר מכך במגזר הציבורי. בתחום החינוך יישום תהליכי מדידה והערכה קשה שבעתיים: תהליכי הלמידה וההוראה מורכבים מאין כמותם, השונות בין התלמידים רבה, קיימות גישות פדגוגיות שונות להשגת יעדי החינוך, תכניות רבות מיושמות במערכת החינוך ומקצת התפוקות והתוצאות מושגות רק לאחר שנים של השקעה.

בחינוך עצמו אין פתרון יחיד העונה על כל הצרכים (one size fits all), וגם אין בנמצא נוסחה אחת ליישום תהליכי המדידה וההערכה שלו. גישות פדגוגיות שונות ותכניות שונות מזמנות, בהתאמה, דגמים שונים של מדידה והערכה. לפיכך, תהליך חינוכי מיטבי צריך להיות מלווה בפעולות של מדידה והערכה שתוצאותיהן יכוונו את העושים במלאכה, וידריכו אותם מה מתאים יותר לתלמידיהם ומה פחות, מה כדאי לשנות ולשפר ומה כדאי לשמר.

הרשות הארצית למדידה והערכה בחינוך (ראמ"ה) הוקמה בשנת 2005 כדי לתת מענה מקצועי לנושאי המדידה וההערכה במערכת החינוך. האידיאולוגיה המונחת בבסיס פעולותיה של ראמ"ה נשענת על שני צירים: א. מדידה בשירות הלמידה; ב. בניית תמהיל של פתרונות מקצועיים, המשלבים באופן אפקטיבי מרכיבים שונים של מדידה והערכה (למידע נוסף על אודות ראמ"ה ראה <http://rama.education.gov.il>).

מבחנים רחבי היקף במערכות חינוכיות ותדירותם

זה כמה עשורים נהוגים במערכות חינוך במדינות שונות בעולם, ובכללן ישראל, מבחנים ארציים המועברים להיקף רחב של תלמידים ומוסדות חינוך (מבחנים רחבי היקף). משמעותם וחשיבותם של מבחנים אלו הולכת וגוברת, והם מטביעים את חותמם על כל השותפים לתהליך הלימודי והחינוכי.

מערכות מבחנים וסקרים מקצועיים רחבי היקף הם כלים חשובים ביותר לניטור ולמעקב אחר הישגי תלמידים ואחר מידת הצלחתה של מערכת החינוך בהקניית ידע וערכים לכלל הלומדים בה. ניתוח אובייקטיבי ומקצועי של תוצאות המבחנים מאפשר לזהות פערים המחייבים תיקון, כמו גם לשלוח זרקור לפינות נשכחות שבהן ראוי להשקיע משאבים גדולים אף יותר. המבחנים עשויים גם

לדרבן למידה, ולהגביר את תחושת האחריות והאחריותיות מצד גורמי המערכת המופקדים על ההוראה. כמו גם, ליצור הלימה רבה יותר בין ההוראה לבין מדיניות משרד החינוך, כפי שהיא באה לידי ביטוי בתכניות הלימודים ובמסגרות ההדרכה וההכשרה של מערכת החינוך.

לצד היתרונות הרבים הגלומים בהפעלתן של מערכות מבחנים אלו, ידוע כי לאורך זמן עלולות להתלוות אליהן גם השלכות שליליות על מערכת החינוך ועל טיב ההליך הפדגוגי בבתי הספר. השלכות שליליות אלו גוברות ככל שהמבחנים נעשים מרכזיים וחשובים יותר בעיני המערכת על דרגיה השונים, וככל שהמבחנים נתפסים בעיני המנהלים, המורים והתלמידים כ"עיתירי סיכון" (high stakes) או חורצי גורל¹. על נטייה זו והשלכותיה כתב פרופ' דון קמפבל (Campbell, 1979), אחד מגדולי החוקרים במדעי החברה:

"The more any quantitative social indicator is used for social decision making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor... achievement tests may well be valuable indicators of general school achievement under conditions of normal teaching aimed at general competence. But when test scores become the goal of the teaching process, they both lose their value as indicators of educational status and distort the educational process in undesirable ways."

ההשלכות השליליות מתועדות זה זמן מה בספרות המחקרית (לדוגמה: Campbell, 1979; Hamilton, 2008; Koretz, 2005; Koretz & Hamilton, 2006; Nichols & Berliner, 2007) ודווחו על ידי ראמ"ה בפרסומים קודמים.

בין ההשלכות השליליות של יישום לא נכון של מבחנים רחבי היקף:

- הסטת משאבי ההוראה מנושאים שאינם נכללים במבחנים, מבתי ספר או מדרגות כיתה שאינם משתתפים בהם, לטובת לימוד נושאי הלימוד הכלולים במבחנים, תוך התמקדות בסוגי השאלות האופייניים להם. תהליך זה פוגע בלימוד ובחשיבות המיוחסת לנושאים אחרים, וכן עלול לפגוע במחזור התלמידים שאינו משתתף במבחנים בשנה נתונה.
- "התכוונות ממוקדת למבחנים" באמצעות למידה אינטנסיבית מכוונת מבחן. למידה כזו מתבססת לעתים קרובות על חזרתיות ושינון, ומערבת פחות מיומנויות חשיבה מסדר גבוה, שהן קריטיות להבנת חומר לימודי ולשליטה בו לטווח ארוך, כמו גם להכללה לתחומי דעת נוספים. נוסף על כך, למידה כזו עלולה לשעמם את התלמידים ולשחוק את חדוות הלמידה, הסקרנות והמוטיבציה שלהם.
- במקרים קיצוניים, מתוך הלחץ שבו נתונים מקצת בתי הספר ומתוך הרצון להעלות את הישגיהם ויהי מה, עלולים אחדים מהם לנקוט פעולות לא לגיטימיות בעליל הפוגעות בטוהר הבחינות (כגון, הרחקת תלמידים חלשים ביום המבחן, ניסיונות להשיג מראש את נושאי המבחן ואת השאלות לצורך הוראתם בכיתה, סיוע לתלמידים בזמן מבחן, ועוד). חמור מזה: פעולות אלו מעבירות מסרים לא חינוכיים לתלמידים.

בד בבד עם הפגיעה האמורה באיכות התהליך הפדגוגי, הפיכת המבחנים לעיתירי סיכון עלולה לפגוע גם בתקפות תוצאותיהם וביכולת להסיק מהן מסקנות שישרתו את המערכת ויאפשרו לה להשתפר. כך, שיפור בהישגים במבחן בעקבות הכנה אינטנסיבית במדגם בתי הספר שבהם הועברו המבחנים בשנה נתונה אינו יכול ללמד על שיפור במערכת החינוך בכללותה, שכן אין הוא מייצג עלייה ברמת

¹ מערכת מבחנים היא "עיתירי סיכון" כאשר מי מן הגורמים במערכת חש מאוים מן התוצאות ותופס עצמו כמי שעלול להיפגע אם התוצאות שלו נמוכות או להרוויח כשהתוצאות שלו גבוהות. עתירות הסיכון גוברות כאשר הפיקוח על בתי הספר נתפס כמאיים על בתי הספר לקראת המבחנים, ונוטה להשתמש בתוצאות כדי לנזוף או להטיל עליהם סנקציות כלשהן.

הידע של כלל התלמידים. שיפור הישגים כזה, גם אם הוא מועיל לכאורה לתדמיתה הציבורית של מערכת החינוך או של חלקים ממנה, הוא במובנים רבים קוסמטי בלבד וחסר ערך בעבור קובעי המדיניות ומקבלי ההחלטות השואפים ליצור שינוי אמיתי ובר-קיימא בכלל המערכת ולאורך זמן. רק תוצאות מבחנים שייאספו ב"תנאי אמת" וללא היערכות מיוחדת לקראתם, יכולות להצביע על מצב המערכת. רק כך יוכלו מקבלי ההחלטות בדרגים השונים ללמוד על נקודות החולשה והחוזק ובהתאם להן לפעול לשיפור המערכת.

על מערכת החינוך, בשיתוף עם ראמ"ה, לפעול נמרצות לצמצום התופעות השליליות הללו, וזאת בראש ובראשונה באמצעות שינוי תרבותי של "מדידה בשירות הלמידה", ולפיו המדידה וההערכה מיועדות לשרת את הלמידה, ולא להפך.

כדי שהמערכת תשתפר עליה להבטיח שכלי המדידה וההערכה שלה יספקו נתונים תקפים ככל האפשר, דהיינו, שתוצאותיהן ישקפו נכונה את מצב המערכת ולא יהיו פרי של מאמץ ייחודי וייעודי שנועד רק להעלות את הציונים במבחן. יש לפעול למיגור התופעות השליליות באמצעות העברת המסרים הנכונים לשדה, ובאמצעות הפחתת הלחץ והאיום של היות תוצאות המבחנים החיצוניים עדות בלעדית לאיכות התהליכים הפדגוגיים בבתי הספר.

עדכון מתכונת המדידה וההערכה: שילוב של הערכה חיצונית ופנימית

דילמה תמידית העומדת לפני מערכת החינוך היא הבחירה בין מדידה פנימית עצמאית של בית הספר (המשוחררת בחלקה מן הלחצים שתוארו לעיל ומותאמת יותר לתלמידים ולחומר הלימוד בכל מוסד חינוך), לבין מדידה חיצונית שהיא סטנדרטית, מקצועית ומרכזית. במילים אחרות, קיים מתח תמידי בין ביזור המדידה וההערכה לבין ריכוזן. יש הגורסים כי הערכה פנימית עצמאית של בתי הספר מתערבת פחות, ומעצימה את מנהלי בתי הספר ואת צוות המורים, בהשוואה להערכה חיצונית. אולם שיקולים של אחריות, אחריותיות, שקיפות, מקצועיות, כלכליות, ובעיקר יכולת השוואה תקפה בין בתי ספר (או בין מגורים, מדינות או קבוצות אחרות), כולל השוואות רב-שנתיות, מחייבים כי חלק מהמדידה יהיה מרכזי וחיצוני, וייעשה על ידי גוף האמון על מדידה והערכה חינוכית.

כדי לשלב בין שתי הגישות, שאף אחת מהן לבדה אינה יכולה לענות על כל הצרכים, וכדי ליהנות מהיתרונות של שתיהן, עודכנה מתכונת ההערכה הארצית בשנת תשס"ז. המתכונת החדשה גובשה על ידי ראמ"ה בשיתוף עם הגורמים השונים במשרד החינוך, בהתייעצות עם מנהלי בתי ספר ועם מורים רבים. היא נועדה לתת מענה מקצועי הולם בתחום המדידה וההערכה החינוכית לכל בעלי העניין במערכת החינוך: לבתי הספר ולגורמים השונים מחוצה להם. המתכונת החדשה נגזרה ממערכת ההערכה שהייתה נהוגה טרם השינוי, תוך ניסיון לשפרה ולתת מענה לחסרונותיה, והיא מבוססת על העקרונות האלה:

- הטמעת תרבות של "מדידה בשירות הלמידה" - מדידה המיועדת לתמוך בשיפור מתמיד של הלמידה, תוך הלימה ליעדים הלימודיים ולחזון הבית-ספרי, וזאת בהתבסס על ההבנה שמבחנים אינם מטרה כשלעצמה אלא כלי בשירות הלמידה.
- שילוב מושכל בין הערכה פנימית להערכה חיצונית ובין הערכה מעצבת (הערכה תוך כדי למידה ולשם למידה) להערכה מסכמת (הערכה של תוצרי הלמידה).
- ביזור מרבי של ההערכה, תוך הקפדה על הסתייעות בכלי הערכה מקצועיים שיסופקו לבתי הספר על ידי ראמ"ה.
- העצמה של מנהלי בתי הספר ושל המורים.

- **הקטנת הלחץ והפחתת התדירות של הבחינות החיצוניות.**
- העדפת השימוש **במבחנים חיצוניים מדגמיים** (מייצגים) על פני מבחנים חיצוניים המקיפים את כלל התלמידים.

המתכונת החדשה שהונהגה בראמ"ה משלבת בין שלושה מרכיבים: הערכה חיצונית מדגמית (בינלאומית וארצית), הערכה בית-ספרית עצמית באמצעות כלים חיצוניים וסטנדרטיים, והערכה בית-ספרית פנימית. כדי לשמור על איזון בין המרכיבים השונים נדרש חיזוק ניכר של ההערכה הבית-ספרית הפנימית, ולשם כך נדרשת הכשרת מורים נרחבת יותר בתחום זה.

המיצ"ב הפנימי (פירוט בהמשך), שהונהג בד בבד עם הקטנת תדירותם של מבחני המיצ"ב החיצוניים, הוא מרכיב חשוב במתכונת החדשה. הוא משרת את מורי בית הספר ואת הנהלתו, ועל-פי חוזר מנכ"ל תוצאותיו אינן מדווחות.

הערכה חיצונית

המיצ"ב

המיצ"ב הוא מערכת של "מדדי יעילות וצמיחה בית-ספרית", הכוללת מבחני הישגים לתלמידים ושאלוני אקלים וסביבה פדגוגית (שמועברים למנהל, למורים ולתלמידים). המערכת מיועדת לשמש למנהלי בתי הספר ולצוותי ההוראה כלי לתכנון משאבים ולניצולם כדי להביא למיצוי הפוטנציאל של תלמידיהם, לשפר את האקלים הפדגוגי ולייעל את מערך ההוראה בבית הספר. ברמה המערכתית מטרות המיצ"ב הינן לספק תמונה על רמת השליטה של תלמידי ישראל בתכנית הלימודים בארבעה מקצועות ליבה, ולשמש את הגורמים המקצועיים במשרד ואת מקבלי ההחלטות במערכת החינוך בקביעת מדיניות-על בסוגיות חינוכיות שונות, כולל אקלים וסביבה פדגוגית.

מבחני ההישגים במיצ"ב מתמקדים בארבעה מקצועות ליבה: **שפת-אם (עברית/ערבית)**, **מתמטיקה, אנגלית ומדע וטכנולוגיה**. דרגות הכיתה הנבחנות בכל ארבעת תחומי הדעת הן ה' ו-ח', ובשפת-אם (עברית וערבית) נערכת בחינה גם בכיתות ב'. מבחני ההישגים נבנים בהלימה מלאה לתכניות הלימודים בתחומי הדעת השונים, ונועדו לבדוק באיזו מידה תלמידי בית הספר היסודי ותלמידי חטיבת הביניים עומדים ברמת הדרישות המצופה מהם על פי תכניות אלו. דוגמאות למבחנים אפשר למצוא באתר ראמ"ה בחלונית "הערכה בית ספרית", בנושא **"מבחני מדף"**.

כל בית ספר משתייך לאחד מארבעה **אשכולות מיצ"ב** – ארבע קבוצות שוות ומייצגות של בתי הספר היסודיים וחטיבות הביניים בארץ (אשכולות א', ב', ג', ד'). כל אשכול של בתי ספר נבחר כך שייצג את כלל בתי הספר בארץ.

מבחני המיצ"ב, במתכונת שנקבעה ע"י ראמ"ה עם הקמתה, מועברים בכל אחד מארבעת תחומי הדעת במחזוריות של ארבע שנים. בתי הספר נבחנים אחת **לשנתיים** במבחנים חיצוניים (**מיצ"ב חיצוני**) בשני תחומי דעת בלבד: מתמטיקה ושפת-אם (עברית/ערבית) **או** אנגלית ומדע וטכנולוגיה, ושנתיים לאחר מכן הם נבחנים בשני תחומי הדעת האחרים. בשנה שבה בית ספר אינו נבחן בבחינה חיצונית בתחום דעת נתון, הוא נבחן בה בתור בחינה פנימית (**מיצ"ב פנימי**) – אותה בחינה המועברת באותה שנה במיצ"ב החיצוני (לפירוט ראה פרק על אודות המיצ"ב הפנימי). כל בית ספר נבחן אפוא בבחינת מיצ"ב חיצוני בכל תחום דעת (בדרגות הכיתה הרלוונטיות) אחת לארבע שנים ובבחינת מיצ"ב פנימי באותו תחום דעת בשלוש השנים האחרות. להרחבה ראו אתר ראמ"ה בחלונית "מיצ"ב", בנושא **"רקע כללי - מיצ"ב חיצוני ופנימי"**.

לוח מס' 1 - מחזוריות אשכולות המיצ"ב לפי תחום דעת ושנת בחינה :

אשכולות בתי הספר	תחום הדעת	תשס"ז	תשס"ח	תשס"ט	תש"ע	תשע"א	תשע"ב	תשע"ג	תשע"ד
א'	מדע וטכ'	חיצוני	פנימי	פנימי	פנימי	חיצוני	פנימי	פנימי	פנימי
	אנגלית	חיצוני	פנימי	פנימי	פנימי	חיצוני	פנימי	פנימי	פנימי
	מתמטיקה	פנימי	פנימי	חיצוני	פנימי	פנימי	פנימי	חיצוני	פנימי
	שפת-אם	פנימי	פנימי	חיצוני	פנימי	פנימי	פנימי	חיצוני	פנימי
ב'	מדע וטכ'	פנימי	חיצוני	פנימי	פנימי	פנימי	חיצוני	פנימי	פנימי
	אנגלית	פנימי	חיצוני	פנימי	פנימי	פנימי	חיצוני	פנימי	פנימי
	מתמטיקה	פנימי	פנימי	חיצוני	פנימי	פנימי	פנימי	חיצוני	חיצוני
	שפת-אם	פנימי	פנימי	חיצוני	פנימי	פנימי	פנימי	חיצוני	חיצוני
ג'	מדע וטכ'	פנימי	פנימי	חיצוני	פנימי	פנימי	פנימי	חיצוני	פנימי
	אנגלית	פנימי	פנימי	חיצוני	פנימי	פנימי	פנימי	חיצוני	פנימי
	מתמטיקה	חיצוני	פנימי	פנימי	חיצוני	פנימי	פנימי	פנימי	פנימי
	שפת-אם	חיצוני	פנימי	פנימי	חיצוני	פנימי	פנימי	פנימי	פנימי
ד'	מדע וטכ'	פנימי	פנימי	פנימי	חיצוני	פנימי	פנימי	פנימי	חיצוני
	אנגלית	פנימי	פנימי	פנימי	חיצוני	פנימי	פנימי	פנימי	חיצוני
	מתמטיקה	פנימי	חיצוני	פנימי	פנימי	חיצוני	פנימי	פנימי	פנימי
	שפת-אם	פנימי	חיצוני	פנימי	פנימי	פנימי	חיצוני	פנימי	פנימי

מחזור 2

מחזור 1

למבחנים הפנימיים נלווים חומרים פדגוגיים לשירות המורים. בית הספר יכול להשתמש בציוני המיצ"ב הפנימי כפי הבנתו, לצרכי שיפור ולמידה פנימיים וכחלק מההערכה השנתית הניתנת לתלמידים. תהליך בדיקתם של מבחני המיצ"ב הפנימי מסייע לפיתוחם המקצועי של המורים, שכן הוא מאפשר להם להיחשף למחווים מקצועיים המגדירים את הציפיות מהתלמידים וללמוד מתוך תשובותיהם על מידת הידע וההבנה שלהם. ציוני המיצ"ב הפנימי משמשים, כאמור, את צוות בית הספר בלבד, ובית הספר אינו נדרש לדווח עליהם לגורם חיצוני (הרחבה על מיצ"ב פנימי בהמשך).

הנורמות הארציות, הנגזרות מתוצאות המבחנים החיצוניים, מדווחות לבתי הספר שעורכים בחינות מיצ"ב פנימי. נורמות אלו מסייעות למנהל בית הספר ולסגל ההוראה לפרש את הנתונים המתקבלים בהעברה הפנימית של אותו מבחן באותה שנה.

סקרי האקלים והסביבה הפדגוגית במיצ"ב נועדו לספק לבית הספר תמונה מפורטת של האווירה בבית הספר ושל התהליכים הפדגוגיים המתרחשים בו, כפי שעולה מהשאלונים לתלמידים ומהראיונות עם המורים. השאלונים מספקים מידע מקיף ורלוונטי על הממדים החשובים בתחום זה, ובהם: רמת המוטיבציה של התלמידים; היחסים בין המורים לתלמידים; אירועי אלימות ותחושת המוגנות של התלמידים; עבודת צוות בקרב סגל ההוראה, ועוד. ממדים אלו מבוססים על תובנות ממספר מקורות: קבוצות מיקוד עם מורים ומנהלים, דיונים עם הגורמים הממונים במטה משרד החינוך, היועצות בחוקרים מהאקדמיה וסקירה של הספרות העדכנית. השאלונים מועברים לתלמידי כיתות ה'-ט' ולמורים בבתי ספר יסודיים ובחטיבות הביניים. בשנת תש"ע הועבר לראשונה שאלון אקלים וסביבה פדגוגית המיועד לתלמידי החטיבה העליונה. בפיתוח מתקדם נמצא עתה סקר אקלים למורים לחטיבה העליונה. להרחבה ראו אתר ראמ"ה בחלונית "מיצ"ב", בנושא ["סקרי אקלים וסביבה פדגוגית"](#).

בסיום תהליך העברת המיצ"ב החיצוני, ראמ"ה מפיקה לבתי הספר דוח מיצ"ב בית-ספרי מקיף (המתייחס להישגים ולאקלים בבית הספר), וכן דוח מיצ"ב ארצי (המתייחס למגזרים, פיקוח וקבוצות חברתיות-כלכליות).

מה אפשר ללמוד מהמיצ"ב

חשיבות השימוש במיצ"ב ככלי עבודה נובעת מהצורך לקבל תמונת מצב דיאגנוסטית עדכנית של רמת ההטמעה והעמידה ביעדים השונים של המערכת (ברמת התלמיד, הכיתה, בית הספר ומערכת החינוך בכללותה), וזאת כדי למצות את הפוטנציאל לשיפור מתמיד של בתי הספר ושל מערכת החינוך.

ברמה הבית-ספרית - תהליך הפקת התובנות מממצאי הדוח הבית-ספרי המפורט מאפשר לצוות של בית הספר לבחון את עצמו ולראות בבית הספר מערכת כוללת (הוליסטית), בהיבטים השונים: הישגים, אקלים וסביבה פדגוגית.

הממצאים מאפשרים לצוות בית הספר לזהות את נקודות החוזק והחולשה בתחומי הדעת הנבדקים, לאתר נושאים או כישורים שלא הושם עליהם די דגש, להעלות השערות בדבר הממצאים שהתקבלו, ללמוד אילו נתונים נוספים יש לאסוף כדי לבסס השערות או להפריכן, לבחון את הסיבות לקשיים שעלו (למשל, מדוע התלמידים מתקשים בכתיבה), ולבנות תכניות ארוכות טווח הנותנות לקשיים מענה מערכתי בית-ספרי. שימוש יעיל בנתוני המיצ"ב יכול לסייע לבית הספר לגבש מנגנונים לשיפור תהליכים בית-ספריים ולתכנן מהלכים ארוכי טווח המשמרים את השינוי לאורך זמן.

ברמה המערכתית - מבחני המיצ"ב החיצוני מספקים ממצאים על מערכת החינוך ועל בתי הספר בחתכים שונים, המסייעים אף הם לזהות קשיים ופערים לצורך שיפור המערכת. כך, למשל, ממצאי ההישגים מושווים על פי הרמה החברתית-כלכלית של בית הספר (מדד טיפוח) ועל פי קבוצות שונות (כגון מגזרים שונים או קבוצות מגדר). השוואות מעין אלו מספקות למקבלי ההחלטות מידע על פערים במערכת החינוכית המחייבים התערבות וטיפול. כן מתאפשרת מאז שנת תשס"ח השוואה תקפה בין הציונים במבחני המיצ"ב בתחום דעת נתון בשנים שונות (ראה להלן), מה שמאפשר את היכולת לנטר את איכות המערכת לאורך שנים. זאת עוד: נתוני המיצ"ב יכולים לשמש להערכת האפקטיביות של תכניות חינוכיות כלל-ארציות המופעלות במערכת. הדוחות השנתיים מתפרסמים מדי שנה בשנה באתר ראמ"ה בחלונית "פרסומי ראמ"ה", בנושא ["נתוני מיצ"ב מערכתיים ורשותיים"](#).

מגמות בהישגים לאורך זמן והשוואת ציונים

שניים מהשימושים הנפוצים, אך לא בהכרח מושכלים, בציוני מבחנים חיצוניים בהיקף ארצי הם השוואות הישגים מעבר לשנים ודירוגים ("טבלאות ליגה"). אולם לא כל השוואה היא תקפה ובעלת ערך; למשל, האם ציון 85 בבחינה קלה השנה מלמד על שיפור לעומת ציון 75 בבחינה קשה אשתקד? ודאי שלא! לכן, כאשר דירוג מעין זה נעשה בין בתי ספר באופן לא מקצועי ובלתי אחראי, הוא עלול לגרום נזק רב.

ראמ"ה פועלת להטמיע בקרב כל השותפים לתהליך החינוכי ובקרב הציבור הרחב את ההבנה שלדירוגים יש משמעות, אם בכלל, רק בקרב אוכלוסיות רלוונטיות, ושל השוואות על פני שנים יש ערך רק כאשר **סולמות הציונים מכוילים**. כיוול נועד לנטרל את ההשפעה של הבדלים ברמות קושי של המבחנים המועברים בשנים שונות באותו תחום דעת, ולאפשר השוואה תקפה בין ההישגים במבחנים אלו. הצורך בכיוול הציונים משנה לשנה נובע מהעובדה שבשנים שונות מועברים נוסחי מבחן שונים, הנבדלים זה מזה ברמת הקושי של השאלות. הציונים **בסולם ציוני הגלם** בכל שנה הם סך הנקודות שצברו נבחנים בתשובותיהם על שאלות בנוסח מבחן מסוים. אי-אפשר לדעת, למשל, אם הסיבה לעלייה בממוצע ציוני הגלם היא שנוסח המבחן היה קל יותר או שהנבחנים הראו רמת ידע גבוהה יותר (או שתייהן).

כדי לפתור בעיה זו ולבחון מגמות בציוני המיצ"ב לאורך זמן, הנהיגה ראמ"ה בתשס"ח מערך כיול סטטיסטי של ציוני המבחנים, המתרגם את הציונים בכל שנה לסולם חדש בר-השוואה - **סולם מיצ"ב רב-שנתי**. הסולם המכיל החדש נועד לאפשר השוואה תקפה של הישגים במבחני המיצ"ב לאורך השנים. מערך הכיול מביא בחשבון את ההבדלים ברמת הקושי של מבחנים שנערכו בשנים שונות, ומאפשר "למקם" את הציונים על פני סולם המדידה החדש. בכך הוא מאפשר לערוך השוואות רב-שנתיות תקפות. סולם המיצ"ב הרב-שנתי נקבע כך **שבשנת הבסיס תשס"ח ממוצע הציונים הוא 500 וסטיית התקן היא 100**.

הליך כיול והנהגת סולם רב-שנתי מקובלים במערכות מבחנים ארציות שונות (כדוגמת המבחן הפסיכומטרי בישראל וה-NAEP בארה"ב) ובכל המבחנים המועברים במסגרת המחקרים הבינלאומיים. עם זאת, יש להדגיש כי ציוני הסולם החדש של המיצ"ב אינם ניתנים לתרגום לציונים במערכות מבחנים אחרות; הם מאפשרים השוואה רב-שנתית **של הישגי המיצ"ב בלבד בתוך כל תחום דעת ודרגת כיתה**.

השוואה של ההישגים במיצ"ב לאורך השנים תשס"ח-תשע"ב מצביעה על **מגמה כללית של שיפור בארבעת מקצועות הליבה** - שפת-אם, מתמטיקה, מדע וטכנולוגיה ואנגלית - בכיתות ה' ובכיתות ח' (למעט מתמטיקה ח'), זאת הן בבתי ספר דוברי עברית והן בבתי ספר דוברי ערבית. בכיתות ה' התאפשרה בדיקה גם בהשוואה לתשס"ז.

- **בכיתות ה'** - בחמש השנים שבין תשס"ז לתשע"ב נרשמה עלייה מצטברת מתונה עד גדולה בארבעת המקצועות. גודל העלייה משתנה בין תחומי הדעת ומגזרי השפה, והוא נע בין 30 ל-60 נקודות בסולם הרב-שנתי.

- **בכיתות ח'** - בארבע השנים שבין תשס"ח לתשע"ב נרשמה עלייה מצטברת בשלושה מקצועות: עלייה קלה באנגלית (כ-10 נקודות) ועלייה גדולה בשפת-אם ובמדע וטכנולוגיה (כ-50 נקודות). הישגי התלמידים במבחן במתמטיקה בתשע"ב דומים להישגים שנרשמו בתשס"ח.

- מגמת העלייה בהישגים לאורך השנים באה לכלל ביטוי בנפרד גם בכל אחת משלוש הקבוצות החברתיות-כלכליות (רקע חברתי-כלכלי נמוך, בינוני וגבוה) בכל ארבעת המקצועות.

- במרבית תחומי הדעת לא נרשמו במהלך הזמן שינויים (צמצום או הרחבה) בפערי ההישגים לטובת תלמידים בעלי רקע חברתי-כלכלי גבוה יותר. יוצאים מכלל זה הם המבחנים בעברית ובאנגלית לכיתות ה': במבחנים אלו הצטמצם בתשע"ב פער ההישגים בקרב תלמידים בבתי ספר דוברי עברית.

המיצ"ב כולל גם בדיקה של **האקלים הבית-ספרי והסביבה הפדגוגית**. בתחום זה הממצאים מצביעים על **יציבות לאורך זמן עם מספר מגמות חיוביות** (בכללן עלייה בתחושת המוגנות, עלייה בדיווח על התנהגות נאותה של התלמידים בכיתה, ירידה קלה בדיווח על אלימות בבית הספר היסודי ועלייה בדיווח על שימוש בתקשוב לצורכי למידה).

הערכה ארצית מדגמית

המחקרים הבינלאומיים

בין שאר תפקידיה, ראמ"ה מופקדת על הפעלת המחקרים הבינלאומיים בישראל. מחקרים אלו מאפשרים להשוות בין הישגי תלמידים ממדינות רבות בכמה תחומי דעת וכן ללמוד על סוגיות חינוכיות אחרות. נוסף על כך, תוצאות המחקרים מאפשרות להשוות בין מגזרים וקבוצות שונות באוכלוסייה בתוך כל מדינה שמשתתפת במחקר. המחקרים מועברים במחזוריות קבועה אחת לכמה שנים, ומאפשרים לבחון מגמות לאורך זמן (גם במחקרים אלו מובנה כיוול). הארגונים הבינלאומיים המפתחים את המבחנים הם מן המובילים בתחום ההערכה והמידה. המבחנים, המתורגמים לשפות השונות, נבנים בקפידה והם בעלי מהימנות ותוקף גבוהים. כל אחד מהמבחנים והשאלונים נבנה לפי מסגרת תיאורטית מפורטת (FRAMEWORK), ששותפים לה מומחים בתחום הדעת ובהוראתו מרחבי העולם.

בכל מדינה המבחנים והשאלונים מועברים למדגם מייצג של האוכלוסייה, והממצאים אינם מדווחים ברמת הכיתה או בית הספר, אלא ברמה הארצית בלבד. ישראל השתתפה בשנים האחרונות בסדרה של מחקרים בינלאומיים, בהם: PISA (אוריינות שפה, מתמטיקה ומדעים לגילאי 15+), TIMSS (מתמטיקה ומדעים לכיתות ח') ו-PIRLS (הבנת נקרא לכיתות ד').

המחקרים הללו מספקים מידע אמיתי ומהימן על מערכת החינוך בישראל תוך השוואה אובייקטיבית למערכות חינוך אחרות בעולם. כמו כן, ההשתתפות במחקרים הבינלאומיים מאפשרת לישראל לעמוד על גישות חדשות ועדכניות בתחומי הדעת הנבדקים, ולבחון את תכניות הלימודים שלה לנוכח הגדרתן במדינות אחרות בעולם. כך, למשל, אותר הצורך לחזק את לימודי האוריינות במדעים, מתמטיקה ושפה, העומדים במוקד מבחן PISA. **המחקרים מאפשרים גם ללמוד על מודלים מוצלחים של מערכות חינוך** במדינות שונות, באמצעות השוואה בין מדינות בעלות הישגים גבוהים לבין מדינות בעלות הישגים נמוכים, בין מדינות השונות זו מזו בפערים בין בתי הספר ובתוכם, ובאמצעות בחינת הקשר בין הישגי התלמידים לבין משתני רקע שונים (כגון השכלת הורים, מצב חברתי-כלכלי, יחס התלמיד למקצוע, ועוד). כך, למשל, למד העולם להכיר את מערכת החינוך המוצלחת של פינלנד, המגיעה להישגים גבוהים ביותר במחקר "פיזה" (PISA), בעיקר בזכות הנורמות התרבותיות והחברתיות הנהוגות בה, שלוו ברפורמה המבוססת על העצמת המורים והמנהלים בבתי הספר.

כלי התקשורת וקובעי מדיניות נוטים לא אחת להדגיש את המיקום במדרג המדינות שאליו הגיעה כל מדינה במחקרים הללו, ולהשוותו לדירוג בבחינות הקודמות, תוך הצגת המיקום הזה כממצא העיקרי העולה מן המחקרים הבינלאומיים. אולם **הערך העיקרי של מחקרים אלו טמון באפשרות לערוך השוואות ובדיקות בתוך כל מדינה בנפרד**, תוך התייחסות להשגת היעדים בכל תחום דעת ולקשר בין ההישגים לבין נתוני רקע ועמדות הקשורות לנושאים לימודיים שונים. בישראל, למשל, תוצאות המחקר מאפשרות להשוות בין הישגי תלמידים בבתי ספר דוברי ערבית לבין הישגי עמיתיהם בבתי ספר דוברי עברית, וכן להשוות בין תת-קבוצות אחרות (בנים ובנות, למשל). בעיקר מאפשרים המחקרים לעמוד על הישגיה של מדינת ישראל לאורך זמן, וזאת בזכות המחזוריות של מחקרים אלו, הנערכים במתכונת דומה אחת לשלוש עד חמש שנים, תוך שמירה על כיוול סולם הציונים ממחקר אחד למשנהו.

בדצמבר 2010 פורסמו תוצאות מחקר פיזה (PISA) 2009 בו נמצא שיפור משמעותי בהישגי תלמידי ישראל בתחום אוריינות הקריאה (ויציבות בביצועי התלמידים באוריינות המתמטית והמדעית). בדצמבר 2013 צפויות להתפרסם תוצאות מחקר פיזה 2012. בעת האחרונה פורסמו תוצאות מחקרי

פירלס 2011 וטימס 2011, שהראו שיפור ניכר בהישגי התלמידים בישראל. בימים אלו אנו נערכים לקראת ההשתתפות במחקר פיזה 2015, שיועבר כולו באופן ממוחשב. פרטים נוספים על כל אחד מהמחקרים האלה אפשר למצוא באתרי המבחנים ובאתר ראמ"ה, תחת החלונית "[מחקרים בינלאומיים](#)".

משוב ארצי מדגמי

ראמ"ה משיקה בימים אלו את פרויקט המשוב הארצי המדגמי שיתמקד בכל פעם בתחום דעת אחר, ויספק מידע ברמה מערכתית על הישגים חינוכיים ועל ההקשר הפדגוגי במערכת החינוך בישראל. המשוב הארצי מאפשר להתייחס לתחומי דעת מגוונים, מעבר לתחומים הנבדקים במיצ"ב. מערכות משוב כגון אלו נהוגות במספר מדינות בעולם; המערכת המפותחת והידועה ביותר היא מערכת המשוב הנהוגה בארה"ב זה עשרות בשנים (NAEP – National Assessment of Educational Progress).

בכל תחום דעת שייבדק במסגרת המשוב הארצי, ייערך המחקר על מדגם מייצג של בתי ספר ויכלול מבחני הישגים באותו תחום ושאלונים המיועדים לתלמידים, למורים ולמנהלי בתי הספר. באמצעות מערך המחקר יהיה אפשר לאסוף מידע על תוצרי למידה (learning outcomes) בתחום הדעת במגוון רחב של תכנים, מיומנויות ואסטרטגיות חשיבה, בקרב **מדגם קטן יחסית של תלמידים**. ריבוי נוסחי מבחן בכל תחום דעת יאפשר כיסוי רחב ומעמיק של התחום ויספק דיווח מהימן על תת-נושאים הכלולים בו. השאלונים יאפשרו לאסוף מידע על אודות משתנים שונים לצורך תיאור ואפיון ההקשר שבו מתנהל תחום הדעת (כגון: פרקטיקות הוראה – למידה – הערכה, פיתוח מקצועי של מורים בתחום, מדיניות בית-ספרית לגבי תחום הדעת), ולצורך מתן הסברים אפשריים לתוצרי הלמידה.

המשוב הארצי יאפשר לקובעי המדיניות ללמוד על הצלחתן של שיטות הוראה וחינוך שונות דרך בחינת הפערים שימצאו בין המגזרים השונים ובחינת מגמות של שינוי לאורך זמן.

לנוכח מטרות המחקר יעובדו ממצאיו וידווחו ברמה המערכתית-ארצית בלבד. לא יעובדו ולא ידווחו נתונים ברמת התלמיד, הכיתה או בית הספר.

המשוב הארצי ייערך במחזוריות של אחת לכמה שנים; כל תחום דעת ייבדק במחזוריות משלו, וזאת כדי לעקוב אחר המגמות המסתמנות לאורך זמן. המשוב הארצי יערך לראשונה בשנה"ל תשע"ד ותחום הדעת הראשון שבו הוא ייערך הוא תחום **הגאוגרפיה**. המשוב הארצי ייערך בכיתות ט', לקראת סוף לימודי החובה בתחום. כלי המחקר (מבחני הישגים ושאלונים) יועברו בסביבה ממוחשבת. המבחנים הממוחשבים בגאוגרפיה יעריכו, בין היתר, מיומנויות גאוגרפיות ע"י שימוש בכלים טכנולוגיים (כגון מפות אינטראקטיביות). באופן זה, ייערך ניסיון לאסוף מידע אודות מיומנויות מסוימות הנדרשות לאזרח במאה ה-21. פרטים נוספים על המשוב הארצי בגאוגרפיה, ובכללם מסמך המסגרת, אפשר למצוא באתר ראמ"ה בחלונית "מבחנים ארציים", בנושא "[משוב ארצי מדגמי](#)". טרם נקבעו תחומי דעת אחרים שבהם ייערך המשוב הארצי.

ניטור ארצי מדגמי של רמת האלימות בבתי הספר

האלימות בבתי הספר היא אחת התופעות המרכזיות שעל סדר היום הציבורי בכל הנוגע למערכת החינוך בישראל. לנוכח העיסוק הרב בסוגיה זו עלה הצורך לנטר את רמת האלימות בבתי הספר כדי לזהות מגמות של שינוי ושימור ברמת האלימות במערכת החינוך.

ראמ"ה ערכה בשנים תשס"ט ותשע"א **סקר ניטור אלימות רחב היקף בקרב מדגם ארצי מייצג של תלמידי כיתות ד'-י"א במערכת החינוך**. הכוונה היא להמשיך ולנטר את רמת האלימות באמצעות

שאלונים דומים שיועברו אחת לשנתיים במדגם מייצג של בתי ספר. הפעימה השלישית של ניטור האלימות מתבצעת בשנה"ל בתשע"ג.

השאלונים שהועברו לתלמידים בדקו שורה של התנהגויות אלימות והתנהגות בסיכון בקרב התלמידים: אלימות קשה, אלימות מתונה, אלימות חברתית, אלימות תוך שימוש באמצעי תקשורת דיגיטליים, אלימות מילולית, חבורות אלימות ומעשי בריונות, אלימות מינית, שימוש בחומרים פסיכו-אקטיביים, אלימות מצד הצוות וכלפי הצוות, הבאת נשק קר לבית הספר, אלימות בהסעות הבית-ספריות, הימנעות מהגעה לבית הספר מחשש לפגיעה, תחושת מוגנות של התלמידים בבית הספר, מאמצי בית הספר במניעת אלימות. פיתוח השאלונים לתלמידים הוא פרי עבודתה של ועדת היגוי ובה אנשי ראמ"ה, גורמים במשרד החינוך העוסקים בהתמודדות עם תופעת האלימות (שפ"י, מינהל חברה ונוער ואגפי הגיל) ומומחים מהאקדמיה.

במרבית מדדי האלימות ניכרת מגמת שיפור בין תשס"ט לתשע"א, בשלבי הגיל השונים ובשני מגזרי השפה. בשאר המדדים נרשמה יציבות בין השנים. בעיקר ניכר השיפור (כלומר ירידה בשיעורי התלמידים המדווחים) במדדים אלו: אלימות קשה; אלימות חברתית; אלימות מצד הצוות וכלפי הצוות; אלימות מינית; אלימות בהסעות; שתיית אלכוהול.

מגמת השיפור מאפיינת את שני מגזרי השפה אך בולטת במיוחד בדיווחי תלמידים בבתי ספר דוברי ערבית, וניכרת בעיקר בשלבי הגיל ד'-ו', ז'-ט'. בדיווחים של תלמידי י'-י"א בבתי ספר דוברי עברית נרשמה על פי רוב יציבות בין השנים. בדיווחי עמיתיהם בכיתות אלו בבתי ספר דוברי ערבית נרשמה על פי רוב מגמת שיפור.

בכל אחת מהשנים תשס"ט ותשע"א: ככל ששלב הגיל גבוה יותר, כך נרשמים פחות דיווחים על התרחשויות של מרבית סוגי האלימות - למעט אלימות כלפי הצוות, הבאת נשק קר לבית הספר, שתיית אלכוהול ושימוש בסמים, תופעות שבהן נרשם פער בכיוון ההפוך.

בכל אחת מהשנים תשס"ט ותשע"א: תלמידים בבתי ספר דוברי ערבית, בהשוואה לעמיתיהם בבתי ספר דוברי עברית, מדווחים בשיעור גבוה יותר על מרבית התופעות השליליות שנבדקו בשאלונים; זאת למעט אלימות מילולית ושימוש באלכוהול, תופעות שבהן נרשם פער בכיוון ההפוך.

פרטים נוספים אפשר למצוא באתר ראמ"ה בחלונית "מחקרים והערכת פרויקטים", בנושא ["ניטור אלימות"](#).

הערכה בית-ספרית פנימית

הערכה פנימית מתקיימת באורח שוטף, מטבע הגדרתה, על ידי צוות בית הספר וביוזמתו. המטרה המרכזית של הערכה פנימית בכיתה ובבית הספר היא לקדם את הלמידה. המידע הנאסף מהערכה מסייע למורה ולהנהלת בית הספר לזהות את נקודות החוזק והחולשה של התלמידים בהתייחס להישגים המצופים בתכנית הלימודים ומנחה אותם בנוגע לצורכי הלומדים ולהתאמת ההוראה לצרכים אלו. גישה זו היא הערכה לשם למידה – (הל"ל), אשר נועדה לתת מענה על שתי שאלות מרכזיות: היכן מצוי הלומד בדרך להשגת יעדי הלמידה? ומהם הצעדים הנדרשים כדי לקדם את הלמידה ולהשיג את יעדיה?

תהליך ההערכה מבוסס על איסוף מידע וראיות ממגוון כלים או מקורות (מבחנים, מאגרי משימות, מטלות הערכה, יחסי גומלין של המורה עם הלומד סביב משימה) ומתן פרשנות אינטגרטיבית לראיות. הפרשנות והסקת המסקנות מהווים תשתית לתכנית התערבות שמטרתה להשיג את יעדי הלמידה.

ראמ"ה מסייעת לחיזוק תהליכי הערכה פנימיים באמצעות בנייה של כלים מקצועיים ושל חומרים פדגוגיים המלווים את הכלים (לדוגמה: מחוונים מפורטים, הסברים על מיפוי השאלות והגדרת הכישורים הנבדקים, מתן דוגמאות לשגיאות נפוצות והצעות לפעילויות הוראה בעקבותיהן, וכדומה).

להלן יוצגו כלים שונים העומדים לרשות המורים לצורכי הערכה בית ספרית פנימית:

מיצ"ב פנימי

ראמ"ה מעמידה לרשות בתי הספר את מבחני המיצ"ב החיצוני, שפותחו באופן מקצועי לצורך העברה פנימית. מבחנים אלו נועדו להשתלב בשגרת ההערכה הפנימית בבתי הספר ולהשלים את כלי ההערכה הפנימיים האחרים המשמשים בהם. בחוזר המנכ"ל שבו דווח על מתכונת ההערכה החדשה נקבע כי ציוני המיצ"ב הפנימי נועדו לשרת את הצוות הבית-ספרי בלבד, ולפיכך אין דרישה לדווח עליהם לשום גורם מחוץ לבית הספר. להרחבה ר' אתר ראמ"ה בחלונית "הערכה בית-ספרית", בנושא "מיצ"ב פנימי".

המיצ"ב הפנימי מושתת על שילוב המרכיבים האלה:

- מבחן ארצי חיצוני אובייקטיבי, בעל איכויות פסיכומטריות של מהימנות ותוקף, שפותח על ידי ראמ"ה בשיתוף ועדות מקצועיות. המבחן משקף את תכנית הלימודים ואת הציפיות הנדרשות מהתלמידים בכל תחום דעת ודרגת כיתה, מבחינת ידע ומיומנויות. בדיקה פנימית על ידי צוות בית הספר (בסיוע מחוון וכלים לצינון) מאפשרת להפיק משוב אישי וקבוצתי מהיר על מידת השליטה של התלמידים בכל תחום דעת.
 - אפשרות להשוואת הישגי התלמידים בבית הספר לנורמות חיצוניות (ארציות, מחוזיות, מגזריות), המתקבלות מעיבוד מבחני המיצ"ב החיצוני.
- התועלות שבית הספר יכול להפיק מהמיצ"ב הפנימי כוללות בין השאר:
- עיבוי של תהליכי ההערכה הפנימיים בבית הספר. בתי הספר יכולים לקבל תמונת מצב בית-ספרית המשלבת מידע שנאסף בכלי הערכה חיצוניים, המאפשרים התאמה להקשר הבית-ספרי. במקרים שבהם לא נעשים שינויים, קיימת אפשרות להשוות את הישגי בית הספר במבחני המיצ"ב לנורמות הארציות.
 - קבלת החלטות מבוססת נתונים. הנהלת בית הספר וצוות המורים יכולים להפיק מתהליך בדיקת המבחנים וממצאיהם, תובנות שיסייעו להם להתמקד ביעדים החינוכיים והלימודיים המתאימים לחזון הבית-ספרי.
 - התאמת בחינות המיצ"ב לצורכי בית הספר. בית הספר יכול להשתמש בציוני המיצ"ב הפנימי לפי הבנתו, כחלק מההערכה השנתית של תלמידיו.
 - צמצום התופעות השליליות שהתלוו לעתים למיצ"ב החיצוני, כגון הסטת משאבים (זמן לימוד וכוח הוראה) על חשבון נושאים אחרים, הרחקת תלמידים חלשים, והיעדר מוטיבציה של מקצת התלמידים להיבחן בבחינה ש"לא הובאה בחשבון" לצורך מתן ציון אישי.

מבחני מדף, מבחנים מעצבים ומאגרי משימות

באתר ראמ"ה (תחת הלשוניות "מבחני מדף" ו"מאגרי משימות") נמצא מגוון רחב של כלים להערכה פנימית בכיתה, ביניהם: שאלונים של מבחני הישגים משנים קודמות, מבדקים ומבחנים מעצבים, מפרטים ומחווני בדיקה ועוד. בנוסף ישנם מאגרי משימות בתחומי הדעת השונים, ובהם משימות

ביצוע המייצגות היבטים שונים של תכנית הלימודים והמיועדות להעריך תהליכים, תוצרים וכישורי למידה מורכבים.

המשימות מיועדות להערכה פנימית במטרה לעצב את תהליכי הלמידה וההוראה. תהליכים אלו מערבים איסוף נתונים ומידע על מגוון רחב של כישורים, ומאפשרים לפרש ראיות ולהסיק מסקנות בנוגע להמשך תהליכי ההוראה והלמידה. המשימות נועדו גם להעריך תהליכים ותוצרים וכן כישורי למידה מורכבים, כגון ראייה מערכתית, פתרון בעיות, נקיטת עמדה, חשיבה ביקורתית, הסקת מסקנות, תכנון תהליכים ומציאת קשרים. המשימות מתייחסות גם לדרך ארגון הלמידה: ביחידים, בזוגות, בקבוצות, ויש משימות שנעשה בהן שילוב של האפשרויות.

להלן רשימה של כלים נוספים המיועדים לתת מענה לצרכים שונים. המידע הנאסף באמצעותם משמש בסיס לפיתוח תכניות התערבות המותאמות לצורכי התלמידים:

מבדק הקריאה והכתיבה בעברית לכיתה א':

- המבדק מיועד לעקוב אחר שליטתם של תלמידי כיתות א' במיומנויות הקריאה והכתיבה בעברית לאורך תהליך רכישת הקריאה.
- המבדק מורכב מסדרת משימות המועברות במהלך שנת הלימודים באופן פרטני (מורה-תלמיד).

מבדק הקריאה והכתיבה בערבית לכיתה א':

- המבדק מיועד לעקוב אחר שליטתם של תלמידי כיתות א' במיומנויות הקריאה והכתיבה בערבית לאורך תהליך רכישת הקריאה.
- המבדק מורכב מסדרת משימות מועברות במהלך שנת הלימודים באופן פרטני (מורה-תלמיד).

ערכה להערכת ראשית קריאה באנגלית לכיתה ה':

בערכה שני מרכיבים מרכזיים: מבדק כיתתי ומבדק פרטני.

המבדק הכיתתי

- נועד לאתר את התלמידים המתקשים בראשית הקריאה על מנת להעביר להם את המבדק הפרטני לצורך זיהוי קשייהם הספציפיים.
- מועבר על ידי המורה לכלל התלמידים.

המבדק הפרטני

- נועד לזהות קשיים ספציפיים המהווים מכשול לרכישת הקריאה.

מבדק עמ"ת בשפת-אם לכיתה ז' (בעברית ובערבית):

מבדק עמ"ת (ערכה למיפוי תלמידים) מורכב משלושה חלקים. בכל חלק יש טקסט אחד ופריטי הבנת הנקרא, לשון וכתיבה. הנבחנים נדרשים לעשות רק שני חלקים מתוך השלושה, על פי הרמה שהושגה בחלק הראשון (חלק א'). מי שעברו את הסף הנדרש ייבחנו לצורך השלמת האבחון בחלק השני (חלק ב'), ואילו מי שלא עמדו בסף הנדרש ייבחנו בטקסט קל יותר (חלק ב').

- המבדק נועד להעריך בראשית השנה את כישורי התלמידים בהבנת הנקרא ובהפקת טקסטים, וזאת כדי לסייע באיתור תלמידים שמתקשים לעמוד בדרישות הלימודים בכיתה ז' ובטיפול בהם.
- המבדק לתלמיד כולל טקסטים מסוגות שונות, שאלות הבדקות איתור מידע ברמת נגישות נמוכה, בינונית וגבוהה, מבנים לשוניים, אוצר מילים וכתיבה.

אוגדן משימות להערכת הבעה בעל-פה בעברית לכיתה ח':

- מקיף שלושה נושאים: קריאה בקול, דיווח צמוד טקסט או ללא טקסט, והשתתפות בדיון קבוצתי.
- כולל משימות הערכה, מחוונים להערכת ביצועי התלמידים ומדריך למורה. כל אלה נועדו להנחות את המורים כיצד להעריך הבעה בעל-פה, והם עשויים להתוות דגם להוראה ולהערכה של הבעה בעל-פה באופן מסודר ושיטתי.
- משימות ההערכה פותחו כחלק מהמיצ"ב הפנימי בשפת-אם עברית לכיתה ח'.

ערכה להערכת השפה הדבורה באנגלית לחטיבת הביניים:

- כוללת חמש יחידות שמשקפות אינטראקציה חברתית ופרזנטציה.
- בכל אחת מן היחידות משימות הערכה, מחוונים למורים ולתלמידים, מדריך למורה, והצעות לפעילויות הוראה שמקדמות את התפתחות השפה הדבורה בקרב תלמידים בחטיבות הביניים.

מבדקים בעברית לתלמידים עולים לכיתות ג'-ט':

- נועדו לבחון מתי ובעזרת איזה סיוע תלמידים עולים יכולים להשתלב בהצלחה בכיתות האם - בקרב בני גילם הדוברים עברית ילדית – ולהשתתף בשיעורים במקצועות הלימוד השונים, תוך כדי קבלת סיוע.
- כוללים חמישה פרקים: שיח, קריאה בקול, הבנת הנשמע, הבנת הנקרא והבעה בכתב.
- מלווים במדריך למורה ובו הנחיות מפורטות מאוד בנוגע למבנה המבדק, מטרותיו, מועדי העברתו, אופן העברתו, משך הזמן הנדרש לשם כך ואופן הציון.

מאגר משימות בתרבות ישראל ומורשתו לכיתות ו'-ח':

- כולל 12 משימות שמשקפות את ארבעת הנושאים המארגנים של המקצוע: היצירה היהודית; מחזור הזמנים: לוח השנה ומעגל החיים היהודי; זיקת עם ישראל לארץ ישראל; דמותה של מדינת ישראל כמדינת העם היהודי. לכל דרגת כיתה ארבע משימות שכל אחת מהן נגזרת מאחד מארבעת הנושאים המארגנים של המקצוע.
- נועד לסייע למורים לפתח ולהעריך את כישורי התלמידים כחלק בלתי נפרד מתהליכי ההוראה-למידה בכיתה. ביצועי התלמידים במשימות אלו ישמשו בסיס לתכנון ושיפור של תהליכי ההוראה, כמו גם למתן משוב יעיל המקדם את הלמידה.
- כל אחת מהמשימות שבמאגר כוללת מידע שיסייע למורה לעשות בהן שימוש מושכל: יעדים (של תוכן ושל רמות חשיבה), הנחיות לביצוע בעבור התלמיד, גריין, ומחווה מפורטת הכוללת תיאור של רמות ביצוע שונות המיועד למורה ולתלמיד, הוראות בדבר דרכי ההעברה של המשימות והנחיות פדגוגיות אחרות למורה.

שימוש במבחני המדף, במשימות ובמבדקים השונים יכול לתרום לשילוב ההערכה בתהליכי הלמידה, לשפר את תכנון ההוראה בכיתה וכן לייעל את תהליך קבלת ההחלטות בכיתה ובבית הספר. ניתוח ממצאי הכלים השונים מאפשר למורים לתכנן פעולות התערבות בהתאם לצרכים השונים של תלמידיהם, כגון: הרכבת קבוצות למידה לפי קשיים שהתגלו, בחירת חומרי למידה מתאימים, חיזוק נושאים שלא נלמדו כהלכה, עמידה על טעויות ועל תפיסות מוטעות נפוצות, ועוד. בתי הספר יכולים גם לקבוע סדרי קדימויות בנוגע להקצאת משאבים ולהתפתחות המקצועית של צוות המורים, בתחומים שתלמידים רבים התקשו בהם.

שימוש בכלים ובאסטרטגיות של מדידה והערכה מעצבת מחייב שינוי תרבותי בבית הספר ובכיתה, לטובת קידום תרבות שבה ההתמקדות אינה בציונים בלבד, אלא גם בתהליך הלמידה עצמו; תרבות שבה נעשה שימוש יעיל במשוב לתלמידים. כמו כן יש לקדם ולהתאים את הסיוע שהתלמידים מקבלים ממוריהם ומעמיתיהם כדי לשפר את למידתם, בהתאם לעולה מתהליכי המדידה וההערכה.

רכז הערכה בית-ספרי

ראמ"ה מייחסת חשיבות מרובה לפיתוח תהליכי הערכה בית-ספריים המסייעים לבית הספר להגדיר את המידע הנדרש לו לתפקוד מיטבי, כמו גם להפיק את המרב מממצאי ההערכה הפנימיים והחיצוניים. כדי לסייע לבתי הספר לאסוף נתונים תקפים ולקבל החלטות על פיהם, פעלה ראמ"ה להגדרת תפקיד חדש במערכת החינוך - **רכז הערכה בית-ספרי**. במסגרת שני ההסכמים "אופק חדש" ו"עוז לתמורה", שנחתמו בשנים האחרונות בין משרד החינוך לבין הסתדרות המורים וארגון המורים, נכלל תפקיד רכז הערכה בית-ספרי ברשימת בעלי התפקידים הזכאים לגמול תפקיד. לפי ההסכמים, בכלל בתי הספר במערכת החינוך ניתן למנות רכז הערכה בית-ספרי אם הוא בעל ניסיון בהוראה ובעל תואר שני במדידה והערכה, או בעל תואר שני בתחום אחר שהשלים התמחות אקדמית במדידה והערכה.

תפקידו של רכז הערכה בית ספרי הוא להוביל להטמעת תרבות הערכה בית-ספרית, בשיתוף בעלי תפקידים אחרים בצוות בית הספר ובכפופות למנהל בית הספר. תרבות הערכה בית-ספרית מניחה כי קהילת בית הספר היא קהילה לומדת, הוראה בהערכה מרכיב מרכזי בתהליך **הוראה-למידה-הערכה** ומשתמשת בכלי הערכה שונים באמצעות מנגנונים התומכים בנורמות ובערכים בית-ספריים, כחלק בלתי נפרד מעבודתה ו/או למידתה. התפקיד כולל עיסוק בנושאי הערכה המשותפים לכלל בתי הספר, אך בעיקר בנושאים ייחודיים העונים על צורכי בית הספר הספציפי והולמים את יעדיו החינוכיים, מאפייניו, השקפת עולמו ותרבותו.

הטמעת תרבות הערכה בית-ספרית מתחילה בבניית מערך ארגוני ומנגנונים בית-ספריים המאפשרים שיתופי פעולה, במקביל לפיתוח ולניהול מאגר כלי הערכה פנימיים ולבניית מסד נתונים בית-ספרי. השאיפה היא שרכז ההערכה יניע שינוי שיטתי ומתמשך בבית הספר, שבמהלכו ילמד צוות בית הספר לקשר בין המטרות הנמדדות בכל אחד מהתחומים או הנושאים הנבדקים לבין תכניות הלימודים. מידע זה יסייע לפרש את ציוני המבחנים בדרך בעלת משמעות, המאפשרת לזהות את נקודות החוזק והחולשה של התלמידים ואת פערי ההישגים בין הלומדים. שימוש במידע זה יסייע לנטר את ההתקדמות ולבחון את השינוי בביצועי התלמידים, וכן לזהות נושאים ותחומים בתכנית הלימודים המצריכים חיזוק, תגבור או שיפור.

רכז ההערכה מנחה צוותים בית-ספריים בכל הנוגע להערכת הישגי התלמידים ולהערכת ההתקדמות בהישגים, תוך שימוש בשיטות הערכה מגוונות וחדשניות ובתוצאות ההערכה לשיפור ההוראה. הרכז פועל להטמעת התפיסה של "מדידה בשירות הלמידה", המדגישה שיפור והתייעלות בדרכי ההוראה והלמידה בבית הספר בעקבות לקחים שהופקו מהנתונים שנמדדו. תפקיד הרכז גם לסייע בפירוש הממצאים בדוחות הערכה בית-ספריים, במאמרים ובנתונים אחרים, ולהפיק מהם השתמעויות ברמת בית הספר שבו הוא עובד. בבתי הספר פועלות תכניות חינוכיות שונות, ורכז ההערכה יכול להוביל תרבות הערכה שבה נאסף מידע שיטתי על תכניות אלו, מה שיוביל לדיון ולהפקת לקחים ברמת מערכתית. לתיאור התפקיד ותכניות ההכשרה ר' אתר ראמ"ה תחת החלונית "הערכה בית ספרית", בנושא ["רכז הערכה בית ספרי"](#).

הערכת סגלי הוראה וחקירת פרקטיקות הוראה

הערכת סגלי הוראה

הערכת מורים² ומנהלים היא נדבך חשוב בתהליך של קידום איכות ההוראה והלמידה (Isore, 2009). זה שנים רבות מנהלים מעריכים מורים ומפקחים מעריכים מנהלים, כל אחד בדרכו: במועדים שונים, בכלים שונים ותוך התייחסות להיבטים שונים.

במסגרת רפורמת "אופק חדש" נקבע סולם חדש לדרגות קידום למורים, לעובדי הוראה אחרים³, לסגני מנהלים ולמנהלים, הכולל גם כמה צמתים שבהם נדרשת הערכה מסכמת. הצורך בהערכה מסכמת כחלק מרצף הקריירה של מורים ומנהלים מזמן, למעשה, הליך הערכה רצוף, מסודר ואחיד לכלל המערכת לצורך קידום מורים, עו"ה אחרים, סגני מנהלים ומנהלים בשלבים שונים.

הכלי להערכת המורים נבנה בשנת תש"ע על ידי ראמ"ה, בשיתוף נציגים של מטה משרד החינוך ומחוזותיו. פיתוח הכלי לווה בקבוצות מיקוד רבות של מפקחים, מנהלים ומורים תוך שימוש בכלים הקיימים בארץ ובעולם. הכלי נבנה כך שישקף את מורכבות עבודתו של המורה וייצור שפה משותפת בין כלל הגורמים במשרד החינוך (מפקחים, מנהלים, מורים ואנשי מטה) בכל הנוגע לביצועי המורה.

הכלי להערכת מורים מבוסס על ארבעה מדדי-על:

מדד-על 1: **תפיסת תפקיד ואתיקה מקצועית**: מתייחס להיבטים הנוגעים להזדהות עם תפקיד ההוראה והחינוך ומחויבות לארגון ולמערכת.

מדד-על 2: **תחום הדעת**: מתייחס לידע בתחום הדעת ולהוראתו.

מדד-על 3: **תהליכים לימודיים וחינוכיים**: מתייחס להיבטים הנוגעים לניהול השיעור וארגונו, לדרכי הוראה, למידה והערכה ולסביבה תומכת למידה.

מדד-על 4: **שותפות בקהילה מקצועית**: מתייחס לשותפות המורה בקהילה המקצועית הבית-ספרית ובקהילה המקצועית בתחום הדעת.

אחד ההישגים החשובים הנגזרים ממהלך פיתוח הכלי להערכת מורים הוא ההסכמה הכלל-מערכתית כי ארבעת מדדי ההצלחה נותנים מענה אחיד ומובנה לשאלה "מיהו המורה הטוב".

הערכת המורים נעשית על ידי המנהלים על סמך איסוף שיטתי של מידע, ובכלל זה תצפיות מתועדות במורה ואיסוף חומרים רלוונטיים אחרים. בנוסף, לצורך ההערכה המורה מתבקש למלא שאלון הערכה עצמית על בסיס אותם מדדי-על. לאחר מכן המנהל והמורה נפגשים לשיחת משוב שמטרתה לדון בפערים בין הערכותיהם. כל תהליך ההערכה נעשה במערכת מקוונת שפותחה במיוחד לשם כך. להרחבה ראו אתר ראמ"ה בחלונית "הערכת עובדי הוראה וניהול", בנושא ["הכלי להערכת מורים"](#) וכן "הערכת עובדי הוראה אחרים".

הכלי להערכת מנהלים נבנה על ידי ראמ"ה על סמך תפיסת תפקיד המנהל שפורסמה על ידי המכון הישראלי למנהיגות בית-ספרית "אבני ראשה", והוא כולל ארבעה מדדי-על:

מדד-על 1: **עיצוב חזון, תמונת עתיד והובלת מדיניות בית-ספרית**. המדד מתייחס להיבטים אלו: עיצוב חזון ממוקד חינוך, הוראה ולמידה בהקשר סביבתי-חברתי; בניית תכנית עבודה בית-ספרית; מעקב אחר יישום התכנית.

² במאמר זה לא נרחיב על מודל הערכת מורים בחטיבה העליונה, כפי שסוכם בהסכם רפורמת "עוז לתמורה", שכן הוא נמצא עדיין בחיתוליו.

³ עו"ה אחרים: גננות, יועצים חינוכיים, עו"ה ממקצועות הבריאות, מתמחים

מדד-על 2: **שיפור ההוראה, הלמידה והחינוך**. המדד מתייחס להיבטים אלו: תכנון הלימודים הבית-ספרי; מיסוד הערכה ולמידה בית-ספרית; תרבות ואקלים בית-ספריים תומכי למידה ולומדים; לגילוי אחריותיות.

מדד-על 3: **הנהגת צוות בית הספר ופיתוחו המקצועי**. המדד מתייחס להיבטים אלו: ניהול ההתפתחות המקצועית של צוות בית הספר; קידום קהילה מקצועית בית-ספרית; איוש, צוות והצמחה של מנהיגות בית-ספרית; אתיקה מקצועית.

מדד-על 4: **קיום קשרי גומלין עם הקהילה**. המדד מתייחס לקיום קשרי גומלין עם קהילת ההורים ועם הקהילה הרחבה.

להרחבה ראו אתר ראמ"ה בחלונית "הערכת עובדי הוראה וניהול", בנושא ["הכלי להערכת מנהלים וסגני מנהלים"](#).

שני הכלים כוללים תיאורים התנהגותיים של מורים/מנהלים בכל אחד מארבעת מדדי-העל, ברמות ביצוע שונות, המייצגות סולם התפתחות מקצועית. הערכת המורה/המנהל נקבעת על פי רמת הביצוע של המוערך בכל אחד מהמרכיבים המפורטים בכלי, תחת כל אחד ממדדי-על אלו. בתור שכאלה, הכלים להערכת מורים ומנהלים משמשים מנוף להתפתחות המקצועית של כלל עובדי ההוראה, ומאפשרים לזהות צרכים ברמת הפרט, בית הספר והמערכת כולה. כלי ההערכה האחרים (הכלי להערכת סגני מנהלים, הכלי להערכת גננות, הכלי להערכת מתמחים, הכלי להערכת יועצים חינוכיים, הכלי להערכת עובדי הוראה ממקצועות הבריאות) פותחו על בסיס אותם עקרונות שתוארו כאן, אך בהתייחס למדדי-על המותאמים לכל אוכלוסייה מוערכת.

סקר מורים בינלאומי

מדינת ישראל הצטרפה למחקר בינלאומי - TALIS – Teaching and Learning International Survey – הנערך במסגרת מנהלת החינוך בארגון ה-OECD. מחקר זה נועד לסייע למקבלי ההחלטות לגבש מדיניות המגדירה תנאים הולמים לקידום הוראה אפקטיבית ובתי ספר אפקטיביים. ההשתתפות במחקר גם מאפשרת ללמוד על הבדלים בין מדינות בתחום המדיניות החינוכית ועל השפעתה של מדיניות זו על סביבת הלמידה של בתי הספר.

המחקר מושתת על מערכת של אינדיקטורים המספקים בסיס מוצק וקוהרנטי של מערכת חינוך "בריאה" ומבוססים על משתנים מדידים של תפקוד מערכות חינוך ושל ביצועיהן. מערכת האינדיקטורים מבוססת על מסגרת מושגית של "הוראה במיטבה" ועל התנאים הבית-ספריים המאפשרים אותה. כך, למשל, הם כוללים איסוף מידע על היבטים אלו: פרקטיקות ההוראה של המורים; האמונות והעמדות שלהם; התפקוד והפעולות של המנהיגות הבית-ספרית; הערכה ומשוב שהמורים מקבלים על עבודתם; האקלים הכיתתי והבית-ספרי; תחושת המסוגלות של המורים ושיעור רצונם מן העבודה.

השתתפות מדינת ישראל במחקר זה מאפשרת למערכת החינוך לגבש ידע ורעיונות בנוגע לפרקטיקות הוראה יעילות ולהטמיע תרבות של למידה ופיתוח מקצועי מבוססי ראיות.

על השימוש בנתוני מבחנים וסקרים לצורכי מחקר והערכת תכניות ופרויקטים

הרצון לשלב את המדידה וההערכה כחלק בלתי נפרד ממערך של תכניות חינוכיות משקף מגמה שמתחזקת בשנים האחרונות, מגמה המייחסת חשיבות לביסוס המחקרי של התערבויות חינוכיות ולממצאים על יעילותן עד כדי הצבתם כתנאי להפעלה. במדיניות שגיבש הממשל האמריקני בראשית שנות האלפיים (שנוסחה בחוק No Child Left Behind), התערבויות חינוכיות נדרשו

להוכיח שהן מבוססות ראיות (evidence based) ושהן נשענות על מחקר קפדני ושיטתי שהניב ממצאים תקפים בטרם יקבלו מימון ליישום. אלא שהדרישה לביסוס מחקרי לא מגדירה בהכרח את תוקף הממצאים. איכות הממצאים תלויה בשיטת המחקר. הממצאים התקפים ביותר מופקים ממחקרים ניסויים מבוקרים שמבוססים על מדגם בהקצאה מקרית (Randomized Control Trials - RCT). מחקרים מסוג זה אינם נפוצים בתחום החינוך, והם מסובכים לביצוע בשל השונות בין בתי הספר, המורכבות של בתי הספר כארגונים, הקושי ביישום אחיד ובשליטה בתנאי הניסוי והבעיות האתיות הכרוכות בעריכת ניסויים כאלה (הענקת תכניות מסוימות לתלמידים מסוימים ומניעתן מאחרים). עקב קשיים אלו יש חוקרים שטוענים כי תחום החינוך שונה מהתחומים הקליניים עד כדי חוסר אפשרות לעשות מחקרים ניסויים בהקצאה מקרית. מנגד, יש הטוענים כי החשיבות של תחום החינוך מחייבת מאמץ מיוחד להפיק מידע מבוסס, שאם לא כן הבחירה במדיניות, בהתערבויות או בפרקטיקות חינוכיות תהיה עניין של טעם או של אופנה הנכפית על בתי הספר בלי שנבחנו העלות והתועלת שלהם.

הוויכוח בין הגישות הולך ומתגבר, אך אין בו כדי לעצור את ההתקדמות בתחום. את השאיפה להפיק מידע מבוסס שיאפשר קבלת החלטות מושכלת אפשר לממש גם באמצעות שיטות מחקר אחרות, זמינות ופשוטות יותר לביצוע, שמספקות ממצאים ברמת תקפות מספקת. מתוך שיטות אלו, השיטה הטובה ביותר היא מחקרים של ניסוי מדומה (Quasi Experiments) המתבססים על קבוצות מושוות (מבחינת המאפיינים). ברמת תקפות נמוכה יותר נמצאים המחקרים שנעשים בשיטה של מבחני קדם-בטר (pre-post); אחריהם - מחקרים שמתבססים על מתאמים ומחקרים שמתבססים על מקרי מבחן (case-studies). בתחתית הרשימה נמצאים מחקרים אנקדוטליים שאינם מאפשרים שום הכללה, ולפיכך ממצאיהם אינם נחשבים תקפים על פי הסטנדרטים המקובלים.

ראמ"ה פועלת מתוך שאיפה להפיק ממצאים תקפים ככל האפשר, ולפיכך היא מתבססת על שיטות המחקר של מחקרים ניסויים (RCT), מחקרים של ניסוי מדומה, מחקרי קדם-בטר, מתאמים ולעתים גם מחקרים על סמך מקרי מבחן. מבחינה זו תיאוריית הפעולה של ראמ"ה קרובה יותר לתפיסה ולפיה מדיניות חינוך צריכה להתבסס על ממצאים תקפים ככל האפשר על אף ייחודו של התחום ולמרות השוני בינו לבין התחומים הקליניים. הדירוג האיכותי של שיטות המחקר השונות בא לידי ביטוי מעשי במאגרי מידע משווים על מחקרי חינוך – מאגרים שמעריכים את איכות הידע הקיים בתחומים שונים של העשייה החינוכית ומסכמים אותו בדרך ביקורתית. אחד המאגרים החשובים הוא What Works Clearinghouse של ה-Institute of Education Sciences במשרד החינוך של ארה"ב. ממאגר זה ושכמותו אפשר ללמוד על תכניות שונות שבנוגע אליהן יש ממצאים תקפים המעידים על מידת יעילותן. המאגרים האלה מעריכים את איכותן של תכניות על פי מחקרים שונים.

הפעילות של ראמ"ה שונה מעט, שכן היא עוסקת בהערכה ולא במחקר תיאורטי של חינוך. הערכה נועדה לספק מידע על תוצאות של תכניות או של מדיניות, בתור אמצעי לשיפור; מידע שיהיה משמעותי למקבלי ההחלטות בזמן אמת (Weiss, 1998). הערכה מתמקדת בממצאים שאפשר ליישם בשדה, ואילו המחקר מבקש להגיע לממצאים שאפשר להכלילם ולקדם בעזרתם את המדע. המחקר מתמקד בתרומה התיאורטית האפשרית ומכך הוא גוזר את שאלות המחקר ואת השיטה, ואילו הערכה גוזרת את השאלות והמטרות מצורכי השדה ומצורכיהם של קובעי מדיניות, ולכן מעמידה בראש את השימושים המיידים, המעשיים והפרטיקולריים. למרות השוני במטרות, שיטות המחקר משותפות, וראמ"ה מתבססת על מערכי הערכה מבוקרים שיצמיחו ממצאים תקפים.

תהליך הערכה של תכנית או פרויקט חינוכי מורכב מכמה אבני דרך: בשלב הראשון נעשה תיאום ציפיות מול מזמין ההערכה, ובו מוגדרות מטרות ההערכה מול מטרות הפרויקט. בשלב השני מוקם צוות מחקר שמתאים לאופי המחקר הנדרש (בדרך כלל שילוב בין מחקרים איכותניים וכמותיים שמופעלים בשני שלבים). בפרויקטים רחבי היקף ממנים גם ועדת היגוי לצד צוות המחקר. צוות

המחקר מגבש הצעת הערכה שכוללת את לוחות הזמנים ואת התקציב, וכן איסוף ידע על התחום באמצעות סקירות ספרות. השלב האחרון כולל את ביצוע ההערכה וכתובת התוצאות בדוח מסכם. הדוח עובר שיפוט, ולאחר תיקון הוא מתפרסם ברבים ומופץ. בסיכום התהליך מתנהל דיון עם הגורמים הרלוונטיים במשרד החינוך לצורך לימוד הממצאים ולשם הפקת לקחים.

התכניות והפרויקטים המוערכים על ידי ראמ"ה מגוונים - הן מבחינת התכנים (חינוכיים, חברתיים, ערכיים, מוסדיים ואחרים) והן מבחינת היקפם ומורכבותם. הפניות להערכה מתקבלות בראש ובראשונה מהנהלת המשרד ומאגפיו השונים (המינהל הפדגוגי, המזכירות הפדגוגית, מחוזות המשרד, המנהלים השונים במשרד). עם זאת, בעת תכנון מערך המחקר נעשה ניסיון לאסוף ידע ומידע שיפנו למעגל רחב יותר מהמזמין הראשוני, והממצאים אף מתפרסמים באתר ראמ"ה.

תהליך ההערכה אינו מבוסס על תכנית קבועה ומוגדרת מראש. קיימת מסגרת תיאורטית שלפיה עוצב מודל להערכת תכניות. מסגרת זו כוללת התייחסות להקשר, לתשומות, לתהליכים ולתפוקות בכל פרויקט שראמ"ה מעריכה (Stufflebeam, 1983), תוך התאמה של דרך המחקר הראויה לפרויקט ולצורכי המזמין. בחינת ההקשר מתחילה בזיהוי הצרכים ובמיפוי של אוכלוסיית היעד ושל הסביבה שבה תופעל התכנית. לאחר מכן לומדים על מטרות התכנית והתאמתה להקשר. בדיקת התשומות כוללת בחינה של הרציונל שעליו נשענת התכנית (בעזרת מומחים בתחום וסקירות ספרות) ושל התאמת התשומות לצרכים. בדיקת התהליך כוללת ניתוח של יישום התכנית מבחינת שיטות היישום, אוכלוסיית היעד, הגורמים המפעילים, ניצול משאבים וכן הלאה (באמצעות סקרים, תצפיות וראיונות). בדיקת התפוקות נעשית באמצעות מדידת האפקט על אוכלוסיית היעד בטווחי זמן שונים והערכה של עלות מול תועלת.

לעתים ההערכה מסייעת למפתחי התכנית לגבש לעצמם את העקרונות, המטרות והתפיסה שעליהן התכנית מתבססת, שכן היא מחייבת להתמודד עם שאלות מובנות על אופי הפרויקט. מדדי התשומות שבהם נעזרים בהערכת הפרויקטים נלקחים ממדדים מינהליים שקיימים במאגר הנתונים של משרד החינוך, ונאספים באמצעות שאלונים וסקרים ייעודיים.

הפעילות של הערכת תכניות חינוכיות ופרויקטים מספקת גם הערכה מעצבת, שאוספת מידע בזמן אמת לצורך שיפור היישום, וגם הערכה מסכמת, שבוחנת את תוצרי הפרויקט. בהערכה מסכמת המטרה המרכזית היא להעריך את האפקטיביות של תכניות או של מדיניות (דוגמת ההערכה שנעשת בתכנית התקשוב הלאומית). הערכת אפקטיביות חוקרת את דרכי היישום ותוצאותיו. אולם יש גם מטרות משנה שחורגות מההבחנה המקובלת בין הערכה מעצבת להערכה מסכמת. כזו, למשל, היא הבקשה לניטור ובקרה של תכניות שונות שמפעילים אגפי המשרד. חשוב להדגיש כי איסוף נתוני ניטור ובקרה אינם בתחום אחריותה של ראמ"ה, אלא הם צריכים להיעשות על ידי מובילי התכניות. עם זאת, נתוני הניטור והבקרה משמשים את ראמ"ה כחלק מהמידע הנדרש לצורך הערכת הפרויקטים.

כלים רבים משמשים את צוותי המחקר: ראיונות מובנים, תצפיות, קבוצות מיקוד, נתוני הישגים (על סמך מבחני המיצ"ב, הבגרות ומבחנים אחרים), נתונים על אקלים בית-ספרי וסביבה פדגוגית, נתונים ממערך ההערכה של סגלי הוראה ומהמבחנים הבינלאומיים. ראמ"ה מתאימה את כלי המחקר המגוונים לצרכים המשתנים של סוגי הפרויקטים השונים ולסוג ההערכה הנדרשת.

כאמור, ראמ"ה מעריכה מגוון רחב של תכניות ופרויקטים מסיבות שונות - קצתן מלוות בהערכה משום רוחב היקפן ובשל חשיבותן למערכת (הערכת רפורמת "אופק חדש", למשל); אחרות מוערכות על אף היקפן המצומצם, וזאת על שום הפוטנציאל שלהן להשפיע על המערכת כולה במקרה של הצלחה (כזו היא התכנית שמבוססת על מודל לחינוך אישי המופעלת בעיר בת ים, או מרכזי השלמת השכלה למבוגרים). תכניות אחרות מלוות בהערכה בשל הימצאותן במוקד התעניינות של הגוף המזמין, למרות היקפן המצומצם (כך, למשל, התכנית למניעת נשירה של האגף לקליטת עלייה או

תכנית אמירים למצוינות). סוג אחר של תכניות מוערך כחלק מהמאמץ לאגם משאבים בראייה מערכתית – לעתים מופעלות כמה תכניות על ידי גורמים שונים, אך המשותף ביניהן עולה על השונה שכן לכולן אותה מטרה, ולעתים אף אותה פדגוגיה ושיטת פעולה (כאלה הן התכניות להעלאת הזכאות לבגרות, התכניות לעידוד קריאה בגיל הגן, וכן הלאה). במקרה כזה ההערכה נועדה למפות את השדה ולתאר את המגוון הקיים, ולבסוף לבחון את האפקט של התכניות במשותף.

ראמ"ה מעניקה עדיפות להערכת תכניות במערך מחקר שניב תוצאות תקפות ככל האפשר. לפיכך, כאשר הדבר אפשרי, נעשה שימוש במערך ניסוי מבוקר. עם זאת, יישומו של מערך כזה כרוך בקשיים רבים כפי שצוין לעיל. דוגמה טובה לכך היא תכנית שילוב טכנולוגיות תקשוב בבתי הספר היסודיים שהפעיל משרד החינוך. לנוכח מגבלת משאבים ביקש המשרד ליישם את התכנית באזורי הפריפריה. ההיגיון הפדגוגי שמאחורי ההחלטה ברור, אך מבחינה מחקרית החלטה כזו אינה מאפשרת לעשות ניסוי מבוקר. למרות זאת נעשה ניסיון ליצור מערך ניסויי שיענה על השאלה האם בתי ספר שמקבלים משאבי תקשוב מתקדמים משפרים את תהליכי ההוראה והלמידה ואת הישגי התלמידים. לקבוצת הניסוי במחקר נדגמו באקראי 60 בתי ספר שהיו שותפים לתכנית, ולקבוצת הביקורת נדגמו 60 בתי ספר שלא היו שותפים לתכנית ושרמת התקשוב בהם הייתה נמוכה. קבוצת הניסוי הוזמנה לכנס ובו הוסבר כי בתי הספר הנוכחים נבחרו למחקר וסופקו להם כלים למימוש התכנית. לבתי הספר שנבחרו לקבוצת הביקורת לא נאמר דבר. הבעיה במחקר התעוררה כמובן כאשר בתי ספר מקבוצת הביקורת התחילו להצטייד באמצעים טכנולוגיים בכוחות עצמם או בעזרת גורמים מחוץ לתכנית (רשויות, למשל), שהרי אי-אפשר לעצור את התקדמותם. מבחינה זו ההבדל בין מחקר קליני ברפואה למחקר בתחום החינוך ברור, שכן במחקר בתחום החינוך אי-אפשר לדמות את השפעתו של כדור הפלצבו ואי-אפשר לשלוט ולפקח על קבוצת הביקורת בדרך הנדרשת בניסוי.

כמו כן יש קשיים בעת ביצוע ההערכה גם בשיטות אחרות: היעדר מסדי נתונים מסודרים בתכניות רבות, פנייה מאוחרת לראמ"ה לשלב מערך הערכה בפרויקט שנמצא בשלבי יישום מתקדמים, המספר הגבוה של מחקרים, מבחנים ושאלונים שבתי הספר מתמודדים אִתָּם, המתח שבין עיבודים מורכבים לבין שמירה על רלוונטיות של הידע המופק מהם, הקושי לאסוף נתונים בעלי משמעות של ממש (לעומת הקלות היחסית של איסוף נתונים שמשמעותם מוגבלת, כגון סקרי שביעות רצון), חוסר הבהירות של מטרות הפרויקט (אפילו למפתחי התכנית), והשינויים הפוליטיים שיוצרים שינויים תכופים בעניין שהמשרד מגלה בפרויקטים שונים.

למרות האתגרים האלה, ראמ"ה עושה מאמץ לשלב מודלים מתקדמים של הערכה ושיטות מתקדמות לעיבוד נתונים, מתוך שאיפה להרחיב את הידע הנאסף בהערכת הפרויקטים ולהפוך אותו למשמעותי אף יותר בתהליכי קבלת ההחלטות.

באתר ראמ"ה, תחת החלונית "[מחקרים והערכת פרויקטים](#)", מוצגים דוחות המחקר וההערכה השונים שנעשו על ידי ראמ"ה.

מדידה בשירות הלמידה - סיכום

שימוש אפקטיבי במדידה והערכה, הן ברמה הבית-ספרית והן ברמה המערכתית, מותנה בקיום מתמשך של תרבות למידה הרואה בממצאי המבחנים והסקרים ובממצאי הערכה פנימיים פוטנציאל לחולל שיפור אמיתי ולא שיפור לכאורה - זו המהות של "מדידה בשירות הלמידה".

הנטייה הטבעית של כל מוערך לנסות לשפר בכל דרך את מיקומו על מדדי ההערכה, כשמדובר בהערכה "עתירת סיכון וסיכוי" בעבורו, קיימת בתחומים רבים ושונים. דוגמאות לכך בתחום הרפואה נוגעות למשל לעובדה שלנוכח הדרישה לפרסם נתוני תמותה בבתי חולים, מסתמנת נטייה

אצל קצתם שלא לקבל אליהם חולים במצב קשה. פרסום מספר הניתוחים לרופא בארה"ב הוביל רופאים בבתי החולים להגדיל בלי הצדקה ובלי פרופורציה את שיעור הניתוחים גם כאשר אין הם נדרשים. באוגוסט 2012 ראה אור פסק הדין של בית המשפט העליון (ע"ס 1245/12), המקבל את דרישת התנועה לחופש המידע לפרסם את ציוני המיצ"ב ברמה הבית-ספרית, ודוחה את עמדת משרד החינוך וראמ"ה ולפיה הנזק הצפוי מהפרסום רב מהתועלת שבו. לפסיקה זו יכולה להיות השפעה משמעותית על מבחני המיצ"ב ואולי גם על עצם קיומם במתכונת הנוכחית. לפירוט ראה אתר ראמ"ה בחלונית "מיצ"ב" – "רקע כללי – מיצ"ב חיצוני ופנימי" בקישור ["מדיניות פרסום תוצאות המיצ"ב"](#).

ראמ"ה רואה את ייעודה העיקרי בסיוע למערכת החינוך להשתפר באמצעות שימוש מושכל במדידה ובהערכה. המושג **"מדידה בשירות למידה"** אינו בבחינת סיסמה בלבד - ראמ"ה רואה בו את ההצדקה לקיומה. ראמ"ה דוגלת בשילוב של הערכה חיצונית והערכה פנימית. שני סוגי ההערכה נקשרים לגישות נבדלות להנעת תהליכי שינוי בארגונים בכלל ובארגונים חינוכיים בפרט. הערכה חיצונית עוסקת בשינוי כולל המתרחש במערכות חינוך ומבוצע ברמה הלאומית על ידי המדינה. גישה זו משקפת שינוי מלמעלה למטה (top down). הערכה פנימית עניינה בשינוי הצומח מלמטה למעלה (bottom up) ומונע ביוזמת המורים או בית הספר כארגון - הם שקובעים את דפוסי הפעולה. הספרות החינוכית ממליצה לנקוט בשילוב יעיל בין שינוי "מלמטה למעלה" לבין שינוי "מלמעלה למטה"; כך גם ראמ"ה מאמינה כי רק דיאלוג ותהליך שפועל בו-בזמן בשני האופנים עשויים לחולל את השינוי המיוחל. יצירת דיאלוג המוביל לקונסנזוס משותף (ברמת בית הספר, ברמת מטה משרד החינוך וברמה החברתית) הוא המסלול המתבקש (Levin, Glaze & Fullan, 2008).

מודל הערכה זה, הכולל הערכה חיצונית רחבת היקף לצד הערכה פנימית מקצועית, שואף להגביל את תפקידה של הערכה רחבת היקף כהערכה קובעת יחידה, ומעמיד לצדה סטנדרטים להערכה פנימית ראויה. הערכה בכיתה עודנה רחוקה מלספק מידע איכותי ולזכות להכרה ולאמון, אבל יש להכיר בעובדה שהערכה רחבת היקף אינה יכולה למלא את הצרכים הללו, ועלינו לכוון את המאמצים ליצירת איזון נאות בין השתיים. לשם כך יש צורך בהכשרה ובהתפתחות מקצועית בתחום ההערכה והמדידה. ככל שתהליכי ההערכה בבתי הספר יהיו מקצועיים יותר, כך תצטמצם השפעתם של המבחנים החיצוניים רחבי ההיקף.

לסיום, מדידה והערכה אפקטיביות מחייבות את קיומם של תנאים אלו: שיתוף פעולה בין כל בעלי העניין, ובינם לבין הגוף המודד; הסכמה על המטרות והיעדים החינוכיים; הסכמה על מטרות המדידה וחידודן; פרשנות והבנה טובות של מערכות המבחנים ושל הממצאים; הנגשת הממצאים ושקיפותם; רציפות ועקיבות בין מחזור מדידה אחד למשנהו; הוגנות כלפי הנבחנים ומוריהם ומיתון ההיבטים המאיימים; מקצועיות - ויושרה.

רק כך תוכל להיות לנו "מדידה בשירות למידה".

- Campbell, D. T. (1979). Assessing the impact of planned social change. *Evaluation and Program Planning*, 2, 67-90.
- Hamilton, L.S. (2008). High-stakes testing. In N. J. Salkind (Ed.), *Encyclopedia of Educational Psychology*, Vol. 1 (pp. 465-470). Thousand Oaks, CA: Sage.
- Isoré, M. (2009). *Teacher Evaluation: Current Practices in OECD Countries and a Literature Review*. OECD Working Papers No. 23, OECD Publishing.
- Koretz, D. (2005). Alignment, high stakes, and the inflation of test scores. In J. Herman & E. Haertel (Eds.), *Uses and misuses of data in accountability testing. Yearbook of the National Society for the Study of Education*, Vol. 104, Part 2, 99-118. Malden, MA: Blackwell Publishing.
- Koretz, D., & Hamilton, L. S. (2006). Testing for accountability in K-12. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 531-578). Westport, CT: Praeger.
- Levin, B., Glaze, A., & Fullan, M. (2008). Results without Rancor or Ranking: Ontario's Success Story, *Phi Delta Kappa*, 90 (4), 273-280.
- Nichols, S. N. & Berliner, D. C. (2007). *Collateral Damage: The effects of high-stakes testing on America's schools*. Cambridge, MA: Harvard Education Press
- Stufflebeam, D.L. (1983). The CIPP Model for Program Evaluation. In G.F. Madaus, M. Scriven, and D.L. Stufflebeam (Eds.), *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*. Boston: Kluwer Nijhof.
- Weiss, C. (1998). *Evaluation: Methods for studying programs and policies*. New-Jersey: Prentice Hall.