

# מדיניות ההכשרה להערכה של סגלי הוראה בישראל: דיון בכמה סוגיות

## מירי לויין-רוזליס

### פרק 6 שער ראשון

גרסה מתקדמת יותר של המאמר פורסמה אצל:  
רמה קלוויר ולאה קוזמינסקי (עורכות) (2012) הבניית זהות מקצועית: תהליכי הכשרה ופיתוח מקצועי של מורים בישראל. תל-אביב: מכון מופ"ת ותירוש הוצאה לאור. עמ' 118-151

## תקציר

הערכה היא חלק בלתי-נפרד ממערכת החינוך בכלל וממעשה ההוראה בפרט. למרות זאת, בכל הוועדות שדנו בהרחבה בתהליכי הכשרת המורים, ובמה שמורים צריכים לדעת, לא דנו כלל בהערכה - לא בהערכה של המורים את התלמידים ולא בהערכה של המורים את עבודת עצמם. זה היה המצב עד שהוקם "כוח המשימה הלאומי לקידום החינוך" המכונה "ועדת דוברת". מסקנות הוועדה עוררו שיח ער ואף מאבקים סביב סוגיית ההערכה, והיו להן שתי השלכות מעשיות: הקמה של רשות ארצית למדידה ולהערכה, והחלטה לשלב רכז הערכה בכל בית ספר בישראל. עם זאת, במסגרות של הכשרת עובדי הוראה אין עדיין חשיבה מערכתית או מדיניות כוללת באשר לתחום הערכה, כך שההערכה נשארת מחוץ לתהליכי ההכשרה של המורה בישראל.

## מדיניות

המונח מדיניות, ובכלל מדיניות חינוכית, משמעו פעולה מכוונת או שינוי מתוכנן בשל אי-נחת מהמצב הקיים (Covaleskie, 1994; Green, 1994; Hall & McGinty, 1997; Spillane, 2004).

### מקומה של ההערכה ביישום מדיניות

אחד ממנגנוני הפיקוח היעילים על יישום של מדיניות או על קידומה הוא הערכה (Datta, 2003; Patton, 1997; Stake, 2007; Weiss, 1991, 1999; Wholey, 1983, 1994). מחקרים רבים בעולם דנים בהשפעה של עצם העובדה שנושא מסוים מוערך, על תשומת הלב שמקבל הנושא, ויותר מכך על שינוי דפוסי ההתנהגות בתחומים המוערכים (Cuban, 2007; Doran, 2003; Greenwald et al., 2006; Howe, 2005; Jennings & Rentner, 2006; Jones & Olkin, 2004; Koretz & Barton, 2004; Messick, 1989). ואכן, בשנים האחרונות משלבים הערכה במקומות רבים שבהם רוצים לשנות דפוסי התנהגות קיימים ולהחדיר דפוסים חדשים. כדי לחזק טענה זו די אם נתבונן בקורה היום בעולם בתחום ההערכה. גופים פיננסיים גדולים וחזקים וגופים אחרים בעלי השפעה עולמית מעורבים היום באופן משמעותי ביותר בשדה ההערכה. לדוגמה, הבנק העולמי מקים ארגוני הערכה ברחבי אפריקה, הממשל הפדרלי בארצות הברית מכתוב סטנדרטים להערכה ולמחקר, האיחוד האירופי ממנה נציגי הערכה במדינות המצטרפות לאיחוד, ממשלות ברוב מדינות אירופה מפעילות יחידות הערכה, האו"ם וסוכנויות שונות שלו מפעילים הערכה ומקיימים הכשרות למעריכים, ה-OECD שותף לארגוני הערכה בינלאומיים חזקים. קמו גופי הערכה רבי-עצמה (כגון NONIE - התארגנות שמאחדת בתוכה 70 ארגונים, ביניהם מחלקת ההערכה של האו"ם, הבנק העולמי, ה-OECD-DAC, וארגונים לאומיים במדינות מתפתחות) שמכתיבים במידה רבה

את שאלות ההערכה ואת דרכי ההערכה בחלק ניכר מגופים אלה (ראו למשל: Chianca, ) (<http://www.oecd.org/oecdworldforum> או <http://www.worldbank.org/ieg/nonie/>). (2004; Cousins & Aubry, 2006; Datta, 2003; Gussman, 2005; Kellaghan, & Greaney, 2001; Levin-Rozalis, Rosenstein and Cousins, 2009; Love & Russon, 2000; Toulemonde, 1999).

מרבית המטרות של גופים אלה, שלעתים מצהירים עליהן חלקית ולעתים אין מצהירים עליהן כלל, הן מטרות של פיקוח, של הטמעת סטנדרטים מתאימים או רצויים לגופים המתערבים, של ביסוס תרבות שלטון שהם יכולים לחיות אתה, של השפעה על התנהלותם של גופים, ארגונים, מערכות ואף ממשלות, של פיקוח על כספים ונהלי הוצאתם, ובמקרים לא מעטים - קולוניאליזם תרבותי.

מערכות החינוך בארץ ובעולם אינן שונות מאלה באשר להערכה כמנגנון פיקוח יעיל. תחום זה, ובמיוחד מבחנים קובעים, משמשים לגופים ריכוזיים לצורך שליטה והכתבה של סדר היום במערכות חינוך ובחברה ככלל (מירו, 2009; Jones & Olkin, 2004; Kellaghan & Greaney, 2001; Kohn, 2000; Lazear, 2005; McDonnell, 2005; Shohami, 2001).

המבחנים הבינלאומיים הם מכשיר פיקוח והשפעה רב-עצמה. הם נתמכים ומבוצעים על ידי ארגוני-על, כמו הארגון לשיתוף פעולה כלכלי - ה-OECD - האחראי למבחני ה-PISA (Program for International Student Assessment), האגודה הבינלאומית להערכת הישגים לימודיים - ה-IEA - האחראית למבחני ה-TIMSS (The International Mathematics and Science Study), הבנק העולמי והאיחוד האירופי, בשיתוף עם גופים כמו אונסק"ו, המקדמים את המבחנים האלה. מטרתם - להשפיע על מעבצי המדיניות במדינות השונות שיכינו את תלמידיהם לכלכלה הגלובלית (Henry, Lingard, Rizvi and Talor, 2001), מבלי שתהיה להם עצמם אחריות לנעשה באותן מדינות (יוגב, לבנה ופניגר, 2009). חוקרים שונים בתחום מתארים את ההשלכות השליליות שיש לדירוג הבין מדינתי על פי ציוני מבחנים (Ertle, 2006; Simola, ) (2005; Romainville, 2002). (Henry, Lingard, Rizvi and Talor, 2001). מבחנים אחרים, מעצם טבעם, אינם רגישים לשוניות. המשמעות המעשית של מבחנים בינלאומיים אחרים היא ויתור על הכרה בשוניות שבין מדינות ובתוכן, ועל בחינת המורכבות התרבותית, החברתית והיסטורית, לטובת הכוונה לגלובליזציה כלכלית (יוגב, לבנה ופניגר, 2009). תפישה זו רואה את סגלי ההוראה כספקי שירותים, וכמי שצריכים לתת דין וחשבון על הצלחתם לספק את הידע הגלובלי הנדרש (Morrow & Torres, 2000).

תחום נוסף שהמדינה רוצה לשלוט בו הוא תכניות הלימודים (Shohami, 2001; Smith, O'Day & Fuhrman, 1994). מדינת ישראל ידעה ויכוחים רבים בנושא, החל בביטול הזרמים בחינוך ועד אכיפת תכניות הליבה. גם תחום זה - תכניות הלימודים - נשלט היטב באמצעות הערכה.

הסיבה לכוח הטמון בהערכה נעוצה כנראה בתכונה האנושית, שאושרה באין-ספור מחקרים ושאפשר לסכמה באופן כללי בכך, שאותם דברים שעל פיהם אנו מוערכים זוכים אצלנו לתשומת לב ולהקצאת זמן ומשאבים (Patton, 1997).

ההשלכות וההשפעות הישירות או העקיפות של עצם השימוש בהערכה ובמדידה, ההשלכות של הערכה בשיטה מסוימת לעומת שיטה אחרת, וכן ההשלכות החברתיות, הארגוניות או האישיות של עצם השימוש בהערכה, במדידה או בבחינה, נקראות "תוקף השלכתי". מושג זה נדון רבות בספרות שעניינה ארגונים ומינהל, מערכות חינוך והערכה (ראו למשל: Cuban, 2007; Greenwald et al, 2006; Doran, 2003; Jennings & Rentner, 2006; Jones & Olkin, 2004; Howe, 2005; Koretz and Barton, 2004;

מתייחס לכמה היבטים אישיים וחברתיים, ואין הכוונה לשימוש בממצאים, ללמידה מהם ולשיפור, או להשפעתם - לטוב ולרע - על המוערכים, אלא להשפעות "מערכתיות" גדולות. בתחום החינוך, ובעיקר בבחינות מסוג מבחנים בינלאומיים או מבחני המיצ"ב, מדובר בעיקר על שלושה ממדי השפעה: (א) הבחינות עשויות להשפיע על התפישות של קובעי המדיניות, של הציבור וגם של המחנכים באשר לשאלה מה הן סוגיות חינוכיות ראויות (ראו מיצ"ב ומבחנים בינ"ל), ובכך הן עשויות לקבוע את סדר היום החינוכי ואף את תכנית הלימודים. (ב) האנשים שמחברים את הבחינות קובעים במובנים רבים את הסטנדרטים החינוכיים. (ג) המערכת תסגל לעצמה נטייה "ללמד/ללמוד לבחינה". כל אלה, מבלי להזכיר את ההונאות הקטנות והגדולות המתפתחות סביב מבחנים ובעיקר גובים מחיר יקר (High stake) (Jones & Olkin, 2004).

### המדיניות כלפי הכשרת עובדי הוראה להערכה

כל חשיבה על מדיניות חינוכית קשורה גם לעובדי ההוראה, ומכאן – להכשרתם. אין חולק על חשיבות איכותו של כוח ההוראה להשגת יעדים חברתיים ותרבותיים שונים, ומחשבה רבה מושקעת בארץ ובעולם בכל הקשור בהכשרתם (Chochran-Smith, 2003; Darling-Hamond, 2000; Elliott, 1999; Freiman-; Nemser & Remillard, 1996; Fullan, 1993; Fullan, 2001; Goodlad, 1995; Richardson, 1997). בקביעת המדיניות של הכשרת עובדי הוראה בישראל שותפים כמה גופים: משרד החינוך, שני ארגוני המורים, המל"ג והמכללות האקדמאיות להכשרת עובדי הוראה. הגוף המעורב והמשפיע ביותר על מערכת החינוך, ובתוך זה גם על המכללות להכשרת עובדי הוראה, הוא משרד החינוך (בריטברד וטביבאן-מזרחי, 2003; נידרלנד, הופמן ודרור, בתהליך), והלה התייחס פעם אחת בלבד לסוגיית ההערכה בהקשר של הכשרת עובדי הוראה, ורק בקשר להסמכתם. "קיום הליך תקין של הערכה הוא תנאי לקבלת רישוי לעיסוק בהוראה" (חוזר מנכ"ל, 2000). הערכה אינה מוזכרת כלל בעניין הידע המקצועי של המורים, אלא כאשר מדובר ברכז הערכה בית-ספרי, שהנו תפקיד מובחן, ולא כחלק מהידע המקצועי של כלל המורים (חוזר מנכ"ל, 2008).

כמה ועדות חשובות, הן של משרד החינוך והן של המועצה להשכלה גבוהה, דנו בסוגיה של עובדי הוראה והכשרתם. בין המוקדמות מזכיר את "הוועדה הציבורית לבעיות הכשרת המורים" (ועדת דושקין, 1962) בראשותו של עמנואל יפה, שהחליטה על רפורמה במבנה המוסדות להכשרת מורים והעמדת משך הלימודים על שלוש שנות לימוד; את ועדת יפה (1971) שטבעה את רעיון האקדמיזציה ובעקבותיה טבע עמנואל יפה את האמירה הדיכוטומית "אקדמיזציה מול הומניזציה" (אורמאן, 1979) המשמשת עד היום, בגלוי או בסמוי, בשיח על הכשרת מורים.

ועדה בולטת אחרת, שהוקמה ב-1979 ודנה בנושא בראשות השופט עציוני, המשיכה לקדם את רעיון האקדמיזציה. המועצה להשכלה גבוהה אישרה את רעיון האקדמיזציה במסמך "הדגם המנחה", שחובר בשנת 1981 ומחייב את המכללות ל"עמידה בסטנדרטים המצדיקים הענקת תואר אקדמי" (ואן-גלדר, 2003; ועדת הקבע למסלולים אקדמיים, 1981). ועדת בן פרץ (2002) קיבלה כבר את האקדמיזציה כמובנת מאליה (נידרלנד, דרור והופמן, 2009; ועדת החינוך של הכנסת, 2002).

בכל הוועדות האלה, שהתייחסו בהרחבה לתהליכי הכשרת המורים, ציינו בכאב את ההישגים הנמוכים - יחסית לציפיות - של הסטודנטים להוראה, והתעמקו בשאלה מה הם צריכים לדעת, לא הוזכרה הערכה מכל

סוג שהוא - לא הערכת התלמידים על ידי המורים ולא הערכת המורים את עבודת עצמם. היכולת להעריך תכניות לימודים או תכניות של התערבות חינוכית לא נדונה כלל, עד שהוקם "כוח המשימה הלאומי לקידום החינוך" בראשות איש העסקים שלמה דוברת ("ועדת דוברת". שאתייחס אליה בהמשך). בהחלטותיה של ועדה זו בולט מקומה של ההערכה - של תלמידים ושל מורים - ואף הוחלט בה למנות רכז הערכה בית-ספרי בכל בית ספר בישראל.

המסקנות של ועדת אריאב (2006) מטעם המועצה להשכלה גבוהה, "מתווה אריאב", יוצאות דופן מאלה של הוועדות שקדמו לה, הן ברוחב היריעה והן במתווה המדויק והמחייב שלה, שמחליף למעשה את ה"דגם המנחה" (משנת 1981). ועדת אריאב שינתה מעט את השיח על המתח שבין "אקדמיזציה" ו"הומניזציה" ומדברת על מורה פרופסיונלי (ועדת אריאב, 2006; נידרלנד, הופמן ודרור, בתהליך). על אף שהוועדה מגדירה בין יתר מרכיבי הפרופסיה, וכבר בפתיחת המסמך, את המורה המקצועי כמורה ש"פועל על בסיס ידע דיסציפלינארי ופדגוגי-דידקטי, באופן שיטתי ומבוסס ראיות, לגבי ההוראה של עצמו והלמידה של תלמידיו" (שם, עמ' 1, סעיף 1) - לימודי הערכה אינם מוזכרים בו. עם זאת, המתווה כולל כחלק מבסיסו (שם, לוח 1, עמ' 6) אוריינות מחקר בחינוך והוראה (לפחות 2 ש"ש), ובהמשך מובהר כי "תחום זה כולל כלים להתפתחות מקצועית של המורה, כגון: קריאת מחקר והבנתו ויכולת לבצע מחקר מורה (למשל: מחקר פעולה, חקר מקרה, מחקר הערכה, סקר)" (שם: 7). עוד מוזכר במתווה, בסעיף קודם שעניינו הוראה דיסציפלינרית, שנוסף לרשימה ארוכה של נושאים שהמורה צריך להכיר בתחום הדיסציפלינה עליו להכיר גם "גישות ושיטות הוראה והערכה המקובלות בדיסציפלינה" (שם: 7). בלימודי התואר השני (M.Teach) מוזכרת מתודולוגיה מחקרית בהיקף של 4 ש"ש (לוח 4, עמ' 12). מתודולוגיה מחקרית איננה כוללת בהכרח לימודי הערכה.

מכיוון שהערכה בכלל, ובחינוך בפרט, היא תחום מורכב ומבלבל, ננסה לעשות קצת סדר בנושא.

## מהות ההערכה בחינוך

הערכה היא חלק מתהליך ההוראה מאז ראשית האנושות (בירנבוים, 1997). מלמדים בכלל ומורים בפרט העריכו תמיד, באופנים שונים, את טיב העבודות של תלמידיהם, אבל להערכה בחינוך יש היבט נוסף, היבט של הערכה מערכתית ושל מדידה, ואלה הולכים ומתמקצעים מאז סוף המאה ה-19 ואילך. ישנם חיבורים, כמו של בינה (שהתבקש על ידי ממשלת צרפת למיין את כלל התלמידים על פי כישוריהם האינטלקטואליים, ובעיקר לחשוף את הבלתי-כשירים), שיחד עם סימון היו הראשונים שניסחו את שאלת הקריטריון לטיב שאלות בחינה, או גלטון (שהתעניין במחוננים), ספירמן (שפיתח כלים סטטיסטיים להשוואת מורכבות של ציוני מבחנים) ופירסון (שפיתח מבחנים סטטיסטיים לבדיקת קשרים בין משתנים, כמו מצב כלכלי וציונים). אלה החדירו למערכת מדידות רחבות היקף (Guba & Lincoln, 1989) הדורשות ידע סטטיסטי נרחב, שלא היה ועדיין איננו מצוי בידי המורה מן השורה. ככל שהלכו תהליכי המדידה החינוכית והתפתחו, כך התרחקו ממהלך ההוראה של המורה בכיתה ומהישגיו המקצועיים. המדידה החינוכית נדדה מן החינוך אל הפסיכולוגיה ואף קיבלה שם חדש - פסיכומטריקה (Levin-Rozalis & Lapidot, 2010).

מאז שטבע רלף טיילור את המושג הערכה (evaluation), כמעט שלא השתנתה הגדרתה הבסיסית - "קביעת ערכו או שוויו של דבר" (Tyler, 1942). מה שהשתנה מאוד הוא היקף היריעה וכמות השאלות

שמעוררת הגדרה פשוטה זו. לא יהיה זה נכון להתייחס להערכה בחינוך מבלי להכיר חלק מהשאלות (לפחות) בה והשפעתן. אפשר שהתפישה הרווחת, שהערכה היא דבר טוב, נכונה בבסיסה, אבל עלינו לשאול את עצמנו שאלות כמו: טוב למי? באילו תנאים? מה המשמעות שלה? מי מרוויח ומי מפסיד, ומה? והשאלות הגדולות הן כמובן מי מעריך?, את מה הוא מעריך וכיצד? לשם מה הוא מעריך? מה יעשה עם התוצאות ואת מי הן תשמשנה? כל שאלה כזו מעוררת שאלות וסוגיות נוספות.

מי מעריך? מעריך חיצוני או פנימי (Love, 1991, 2005; Nevo, 2001)? המושג "פנימי-חיצוני" מבחין באופן כללי בין הערכה הנעשית על ידי הישות המוערכת בעבור עצמה, ובין הערכה הנעשית על ידי גורם חיצוני לצרכיו הוא. היא נוגעת ברמות שונות, החל ברמת הכיתה ובהערכת הישגי התלמידים, וכלה ברמת המוסד או אף המערכת (החל בתלמיד שמעריך את עצמו וכלה במבחנים בינלאומיים, חיצוניים למדינה).. מומחה מתחום התוכן, שהערכה איננה בהכרח תחום מקצועי שלו, או מומחה בהערכה שהידע שלו בתחום התוכן עשוי להיות מוגבל (Levin-Rozalis, 2003)? מכאן נובעות שאלות נוספות, כמו איזו הכשרה נדרשת למעריך? מהו הסטטוס ומהו התפקיד של המעריך, ושל ההערכה בתוך בית הספר? מי מחליט על כך (ליון-רוזליס ולפידות, 2009)?

את מה או את מי מעריכים: את תכניות ההתערבות? את המוסד החינוכי? את המורה/ המנהל/ התלמיד? את תכנית הלימודים? את המדיניות החינוכית/ החזון החינוכי של המוסד? את הישגי התלמידים? כיצד מעריכים? האם מעריכים תהליך או תוצר (Stake, 1967)? האם ההערכה היא מעצבת ונעשית כדי לשפר את המצב הקיים, או מסכמת לשם קבלת החלטה (Scriven, 1967)? באילו כלים מעריכים (Cronbach, 1983)? לשם מה מעריכים? למה ואת מי תשמשנה התוצאות? האם ההערכה היא לצורך למידה ושיפור, או לצורך פיקוח ובקרה ואפילו הענשה?

השאלות הללו נובעות מתפקידיה השונים ואף הסותרים של ההערכה בחינוך, שמקורם בשתי תפישות מרכזיות - זו המציבה את המבנה במרכז לעומת זו המציבה במרכז את האדם הפועל, את הסוכן האנושי (Berger & Luckmann, 1966). על פי תפישת המבנה, המערכת החברתית והמבנה שלה הם המשפיעים על המציאות של היחידים, ורק אם נבין את המבנה נוכל להבין תופעות חברתיות שונות. ואילו גישת הסוכן האנושי או הפועל, תופשת את היחידים כמשפיעים על עולמם, ורק אם נבין את התהליכים של היחיד נוכל להשפיע. אפשר לארגן את תפקידי ההערכה בחינוך בדיכוטומיה הזו באופן המוצג בלוח הבא (Levin-Rozalis, ;Rosenstein & Cousins, 2009 התרגום שלי, מל"ר).

### לוח מס' 1: תפישות הערכה סותרות והשלכותיהן על דרכי הערכה

הממד	גישת ה"מבנה"	גישת "הסוכן האנושי"
הפונקציה	פיקוח, בקרה ושליטה	למידה
המטרה	סטנדרטיזציה	הכרה בשונות, במגוון, בהבדלים, ובחינתם
המסגרת	המבנה. מבט מערכתי	דיאגנוסטית (הבחנה בין תלמידים, מורים, או מוסדות)
המוקד	תוצרים	תהליכים (למשל, הערכה חלופית)
התוצאה	מיון	חיזוק, פיתוח וקידום

התוצר החינוכי	ידע	מיומנויות
המתודולוגיה	ניסויית	מגיבה
גישת החקר	אנליטית	הוליסטית
המיקוד	חיצוני	פנימי

מעניין לראות את ההקבלה (שאיננה מוחלטת) בין התפישות הפסיכומטריות ובין גישת "המבנה" (הטור הימני), ואת זו שבין גישת "הסוכן האנושי" (הטור השמאלי) ובין עבודתו של המורה בכיתה. נציין שאין כאן בהכרח הקבלה בין גישות כמותיות או איכותניות לאיסוף נתונים. אין זה דיון אקדמי על הגדרה זו או אחרת של הערכה, אלא על ההשלכות המעשיות שלה. על הדרך שבה נעשית ההערכה, על שאלותיה, פְּלִיָּה, אמות המידה והסטנדרטים שלה, על יחסי הכוח שבין המערך למוערך, שבין יוזם ההערכה למוערך, ועל יחסי האמון ביניהם. לכל אלה יש השפעות מרחיקות לכת על הדרך שבה פועל ויפעל המוּעָרֵךְ (מושא ההערכה), בין אם זה ארגון, מורה, תלמיד או מערכת. לכל בחירה באשר לדפוס ההערכה ואופייה במערכת החינוך יש השפעה רבה על אופי מערכת החינוך, ובעיקר על דפוס ההוראה והלמידה (בירנבוים, 2006; לוי, ללא תאריך; לוין-רוזליס ולפידות, 2010; Bloom, Alkin, 1970; Levin-Rozalis & Lapidot, 2010).

### המתחים המרכזיים לגבי תפקידי ההערכה ומשמעותם למערכת החינוך

בבסיס התפישה של תפקידי ההערכה קיימים כמה מתחים:

- א. מתח בין הערכה כתהליך של בקרה ופיקוח ובין הערכה לצורך למידה
- ב, מתח בין הערכה לפי מודלים מבניים קשיחים ובין הערכה מגיבה ותלוית-שדה
- ג. מתח בין הערכה של תוצרים ובין הערכה של תהליכים ומורכבויות.

נבהיר מעט:

#### א. מתח בין הערכה כתהליך של בקרה ופיקוח ובין הערכה לצורך למידה

הערכה משמשת רבות לצורך פיקוח. אחד המושגים המשמשים היום להטבעת מערכות פיקוח במערכת החינוך הוא "אחריותיות" (accountability). המקור של המושג ושל התפישה שברקע שלו הוא תאוריות וגישות של ניהול מדרגי (היררכי), לפיהן הכפוף מוסר דין וחשבון לממונה עליו. מושג זה עבר גם להערכה, בעיקר של הזרם שדורש שינוי במערכת החינוך על פי המטפורה של "מפעל": תשומות ותפוקות מדידות, ברורות וחד-משמעיות, עם קשר חד-חד ערכי ביניהן, ללא גורמים מתערבים, ללא השפעות ארוכות טווח וללא תהליכי ביניים.

הרעיון של אחריותיות קיבל משנה תוקף בארה"ב, עם אישור החוק של "No Child Left Behind" (NCLB) – The reauthorization of the Elementary and Secondary Education Act 2001.

מאז צבר הנושא תאוצה, וכך גם הוויכוח בעקבותיו (Hamilton, Stecher & Klein, 2002). אחת הבעיות המרכזיות בו היא אי-הבהירות של המושג, הבלבול בינו ובין "אחריות", וההבנות השונות של

מקבלי החלטות שונים (Darling-Hammond & Snyder, 1992; Heim, 1995). פרידמן (2005) מונה שש דרגות של ביטויי אחריותיות שההשלכות שלהן הולכות ומחמירות:

1. דיווח על ביצוע באמצעות ממצאים מדידים. הבעיה: תהליכים, מורכבויות ותופעות שאינם מדידים אינם נכללים בו.
2. פיקוח ועמידה בסטנדרטים – ביצעת או לא ביצעת, זו חזות הכול. הבעיה: הרלוונטיות של הסטנדרטים, נטייה של אנשים לא לקחת סיכונים, לא להסתכן בעשייה שסיכויי ההצלחה שלה נמוכים (עבודה עם תלמידים מתקשים, למשל).
3. תמריצים – תגמול לפי התוצאות (מחייב לעמוד בשלבים הקודמים) – תגמול דיפרנציאלי על פי מידת העמידה בסטנדרטים.
4. כוחות השוק. בשלב זה מתערבת העיתונות, ומתחילים לחצים פוליטיים שונים, רלוונטיים ולא-רלוונטיים, לעתים מתוך אינטרסים זרים ופעמים רבות מתוך אי-הבנה של העניין. כתוצאה מהלחצים האלה נקבעים סטנדרטים פופוליסטיים שלעתים אף מזיקים,
5. בקרה חיצונית והקטנה של אחריות אישית. כתוצאה ממערכת הלחצים מתחמקים המוערכים מאחריות אישית לביצוע מקצועי ועושים הכול כדי להתאים את עצמם לסטנדרטים. אלה שאינם מוכנים לכך - נפלטים מהמערכת.
6. הגברת הלחץ והפיקוח הפנימי –אווירה לחוצה בתוך הארגונים. יש לחץ של כולם על כולם לעמוד בסטנדרטים, מתחילה תחרות ונעלם שיתוף פעולה שאינו מתוגמל ואף נענשים עליו.

מנגד, מדובר בהערכה כמשרתת למידה - למידה של תלמיד או של מורה, של מוסד ואף של המערכת כולה. ההבחנה בין הערכה המשמשת ללמידה ובין הערכה המשמשת לפיקוח היא באמצעות "מבחן הנתונים". את מי אמורים הנתונים לשרת: את מושא ההערכה (תלמיד, מורה, מוסד, תכנית, מערכת) או את מי שמחוץ למושא ההערכה ויש לו השפעה על גורלו? (מורה מקבל את נתוני התלמידים, מנהל מקבל את נתוני המורים, וכן הלאה.) אם מזמין (ומקבל) הנתונים איננו מושא ההערכה אלא מי שמחוץ לו - מדובר למעשה בבקרה, גם אם בחלק מהמקרים מושא ההערכה מקבל אף הוא את הנתונים.

למידה אמיתית מתרחשת בעיקר אם שאלות ההערכה מנוסחות בשיתוף מושא ההערכה ומותאמות לצרכיו. בדרך זו הוא לומד לשאול את עצמו שאלות על עצמו, על תפקודו, על הישגיו ועל יכולתו להגדיר מטרות ולפעול ליישומן. לעומת זאת, אם השאלות מונחתות עליו, - הלמידה תהיה חלקית, אם בכלל, וחלק ממנה יהיה כיצד להתחמק או כיצד "לצאת טוב" על פי אמות המידה (קריטריונים) שקבע מישהו אחר, ולא כיצד לשפר את הטעון שיפור.

## **ב. מתח בין הערכה במודלים מבניים קשיחים ובין הערכה מגיבה ותלויה שדה**

במודלים מבניים קשיחים, הערכה משמעה מבחנים שנועדו לתת תמונה מערכתית של המוערך - כיתה, בית ספר או כלל המערכת (באמצעות מבחנים רחבי היקף או מדגמיים. מדובר בבדיקה שרמת אמינותה גבוהה, וחשוב מכך - בעלת יכולת הכללה גבוהה. אחד המודלים המבניים הקיצוניים הוא המודל הניסויי (הקצאה אקראית של המוערכים לקבוצת ניסוי וקבוצת ביקורת – RCT: randomized controlled trials),

שלמרות יתרונותיו הוא שנוי במחלוקת קשה, מכיוון שאופן הבדיקה מכתוב את אופן ההתערבות. תהליך הבדיקה חשוב יותר מאשר שיקולים מקצועיים (איך להקצות) או אתיים (קיפוח אלה שלא הוקצה להם). לא רק למודל הניסויי אלא לכל המודלים המבניים הקשיחים יש חסרונות, שעיקרם העדר רגישות לתהליכים ולשנוניות.

בקצה השני נמצאות המדידה וההערכה המגיבות, תלויות השדה. אלה בודקות את המוערך באופן תהליכי, כחלק מההקשר שבו הוא פועל ותוך שיתוף פעולה ודיאלוג בין המעריכים למוערכים. כלומר, הבדיקה מתאימה את עצמה, את כלי הבדיקה, את שיטות הבדיקה, ואת שיטות ניתוח הנתונים לתנאי השדה, להתערבות, לאוכלוסייה וכדומה. זוהי בדיקה המותאמת למידות המוערך.

למען הסר ספק ולמנוע טעות, בדיקה מגיבה איננה בהכרח איכותנית. אפשר לעשותה באופן כמותי, באמצעות שאלונים, מבחנים או כלי מדידה אחרים, והממצאים מעובדים ומנותחים בשיטות סטטיסטיות מוקפדות. כמו כן, הערכה במערך ניסויי איננה בהכרח כמותית: אפשר להפעיל בה כלי פתוחים כמו ראיונות, תצפיות או ניתוח תוכן של מסמכים/מבחנים, אולם העיבוד הסופי ייעשה בסופו של התהליך במספרים ובכלים סטטיסטיים, על מנת לקבוע את מובהקות ההבדלים.

דפוס הבדיקה שישלט במערכת החינוך יכתיב במידה רבה את אופי הפעולה החינוכית, את אופי ההתערבויות החינוכיות ואת המבנה שלהן, ובסופו של דבר גם את המשתנים שיהוו את מושאי ההערכה ואת אופיים (האם ייכנסו למשוואות רק משתנים מדידים, ואז יהיה צורך לכמת כל דבר, או שתופעות ותהליכים מקומיים, ייחודיים, הוליסטיים, נטורליסטיים ועמומים יותר ימלאו תפקיד בהבנת השדה החינוכי).

יש גורם נוסף המבחין בין שני הקצוות שהוצגו כאן, וזו סוגיית המקצועיות של המערך. המודלים המבניים הקשיחים מצריכים רמת מקצועיות גבוהה בכל הקשור בבניית הכלים, בתכנון של הבדיקה, המדגמים והניתוחים הסטטיסטיים ובהבנת משמעותם. מורים אינם יכולים לבצע את פעולות ההערכה האלה, אבל עליהם להיות מסוגלים לקרוא דו"חות הערכה כאלה ולהבינם, בעיקר אם התפישה המובילה היא של הערכה ומדידה לצורך למידה. מורים אינם יכולים לבצע גם את כל אמור בהערכה מגיבה, אבל הם יכולים להיות שותפים לדיאלוג בנושא מה מעריכים ומדוע ומה הן שאלות ההערכה או המדידה, ולמסור למעריכים מידע רב הקשור במוערך (לויין-רוזליס, 2005). מורים יכולים לבצע, בהנחיה ואף בלעדיה, הערכה של תלמידים בגישות חלופיות, הערכה עצמית והערכת עמיתים, בתנאי שנושאים אלה יהיו חלק מהכשרתם (הוועדה למדידה ולהערכה בחינוך, 2005; הרטף וחוב' 2007).

### **ג. מתח בין הערכה המדגישה תוצרים בלבד ובין מדידה והערכה המתייחסות גם לתהליכים**

הערכה של תוצרים הנה חשובה, מכיוון שהיא מאפשרת לדעת את מצבם (הישגים, חוסרים, צרכים וכו') של חלקים שונים במערכת. אי-אפשר להימנע מבדיקה כזו, וחשוב שתיעשה בצורה מקצועית ואמינה, החל במורה בכיתה. הערכת תוצרים עונה על השאלה "מה" - מה קיים ומה הושג. וכדי לענות על "למה" ו"איך" יש להתבונן בתהליכים ובמורכבויות. במסגרת ההערכה של תהליך אפשר לכלול את נושא ההערכה החלופית בהערכת הישגים.

יש קשר בין שלושת סוגי המתח האלה. בדרך כלל, הערכה לצורך פיקוח ובקרה בוחנת בעיקר תוצרים במערכי בדיקה נוקשים. נכון שאפשר גם וגם, כלומר אפשר לשלב. אולם יש לזכור שמאחורי כל צד במתחים אלה מסתתרת תפישת עולם מחקרית-מתודולוגית, אפיסטמולוגית, והרבה פעמים אף ערכית ואידאולוגית. החיבור של גם וגם איננו פשוט, לא תמיד אפשרי, ובמקרים רבים פשוט לא נעשה.



## תפישות מרכזיות בשיח על הערכה חינוכית בישראל

בהשוואה לעולם המערבי, ההערכה בישראל בתחום החינוך ומחוצה לו התפתחה בעצלתיים. תכניות ההכשרה היו מעטות, ופעולות ההערכה שנעשו, גם רחבות ההיקף שבהן, היו בעיקרן יוזמות של יחידים או דרישות של קרנות חינוכיות (הלוי, 2009; Schwartz, 1998; Levin-Rozalis & Shochot-Reich, 2008). עבודות הערכה של משרד החינוך היו אקראיות, לא סדירות, והייתה בהן פעולה משמעותית אחת בלבד **הקשורה בהכשרה של סגלי הוראה**. ב-1993 חברו יחד משרד המדען הראשי, האגף לחינוך יסודי והאגף לתכניות לימודים כדי לפתח מאגרי משימות להערכת לומדים. מאגרי המשימות של "המשוב הבית ספרי", שפותחו בגישה חדשה להערכה (הערכת ביצוע), נועדו לשפר את איכות ההוראה-למידה-הערכה בבתי הספר. הם פותחו על ידי מומחים בחמישה תחומי תוכן בעבור שתי דרגות כיתה עיקריות - ד', ו' - אך שימשו בפועל ארבע דרגות כיתה - ג'ו'. במשרד הוקמה יחידת "המשוב הבית ספרי", שנועדה להטמיע את מאגרי המשימות ולעודד בתי ספר לבצע הערכה פנימית. עם כניסת המיצ"ב דרך השימוש במשימות, והמאגר נמצא באתר הראמ"ה (הרשות הארצית למדידה והערכה) לשימושם של המורים (מירו, 2009).

עם פרסום ההמלצות של כוח המשימה הלאומי לשיפור החינוך (התכנית הלאומית לחינוך, 2005), השתנתה באחת ההתייחסות להערכה. כוח המשימה הלאומי המליץ לשלב איש הערכה בכל בית ספר בישראל. המלצה נוספת של כוח המשימה הלאומי הייתה להקים גוף הערכה בלתי-תלוי – רשות ארצית למדידה והערכה (ראמ"ה). לשתי ההחלטות יש השלכות מרחיקות לכת על תפקיד המורה ועל מערכת החינוך. החלטות אלה העלו את נושא ההערכה בכלל, ואת ההערכה החינוכית בפרט, על סדר היום הציבורי. היום מתקיים דיון ציבורי נרחב בנושא הערכה בחינוך ושותפים לו - באמירה או במעשה - גופים שונים: האקדמיה הלאומית הישראלית למדעים (באמצעות הוועדה למחקר יישומי בחינוך ובשיתוף יד הנדיב), המועצה להשכלה גבוהה, משרד החינוך, גופים המכשירים פרחי הוראה (מכללות ואוניברסיטאות), גופים להשתלמות מורים (מכללות, מרכזי פסג"ה<sup>1</sup>, אוניברסיטאות, גופים פרטיים ואחרים), איל"ת (האגודה הישראלית להערכת תכניות), אפ"י (אגודה פסיכומטרית ישראלית) ועוד. לכל אחד מהגופים יש תפישה משלו על הערכה, סדר יום משלו, וכל אחד מהם מושך לכיוון שלו, רק המדיניות להכשרת עובדי הוראה להערכה לא השתנתה. השיח הער והמשמעותי כמו פסח עליה, אם כי גופים מסוימים בעלי השפעה הכניסו את המורים לשיח שהחל להתפתח בנושא, כפי שאבהיר בהמשך.

ננסה לבחון את התפישות של כמה מן הגופים השותפים בשיח, את פעולותיהם בעניין זה ואת משמעות התפישות והפעולות להוראה כמקצוע, ומכאן להכשרת עובדי הוראה בישראל.

### כוח המשימה הלאומי לשיפור החינוך (ועדת דוברת):

התפישה שבאה לידי ביטוי בהמלצות של כוח המשימה הלאומי (התכנית הלאומית לחינוך, 2005) היא בעיקרה ניהול מבוסס תוצאות (evidence based management), שמשמעותו אחריותיות (accountability) ובקרה. כך הוגדרה הבעיה על ידי הוועדה: "...תרבות קבלת החלטות אינה נשענת באופן הנדרש על קבלת נתונים וניתוחם המקצועי... תכניות לימודים רבות אומצו בעבר על סמך עמדות (ואופנות), ולא התבצעו תהליכים מסודרים של מדידה והערכה לבדיקת יעילותם..." (שם: 29).

<sup>1</sup> פסג"ה – פיתוח סגלי הוראה, גף במשרד החינוך שאחראי לנושא של השתלמויות מורים מכהנים. מרכזי הפסג"ה מפוזרים בכל רחבי הארץ.

וכך גם בהגדרת הפתרון: "...אנו רואים חשיבות מיוחדת בפיתוח תרבות ארגונית חדשה המקיימת יישום מלא של עקרונות האחריותיות (Accountability)... הערכה ומדידה מתמידות של הפעולות והתוצרים, כמו גם הקפדה על שקיפות ופרסום מלא של נתונים, ובכלל זה נתונים לא מחמיאים, הם תנאים הכרחיים למימוש עקרונות האחריותיות בכל דרגיה של מערכת החינוך..." (שם: 37). בהמשך מסבירה הוועדה את המשמעות של אחריותיות: "...על-פי עקרונות האחריותיות (Accountability), כל דרג חופשי מהתערבות מיותרת בפעולתו מצד הדרג הממונה עליו, אולם חייב במדידה ובהערכה של פעולותיו והישגיו תוך שקיפות מלאה... משנקבעו היעדים, יש לבקר ולוודא כי הפעילות נעשית בהתאם ליעדים ולכללים שנקבעו..." (שם: 42).

ובמקום אחר: "...בהקשר זה נודעת חשיבות מיוחדת למדידה ולהערכה. כלי מדידה אמינים והערכה מקצועית הם קודם כול האמצעי המשמעותי ביותר למורים ולמנהלים בבתי-הספר לניהול תהליכים חינוכיים מכווני תוצאות. זהו גם האמצעי העומד לרשות משרד החינוך בבקרת פעולתם של מינהלי החינוך האזוריים, כאשר אלו מצדם יבצעו הערכה מקצועית של בתי-הספר הכפופים להם" (שם: 42). ההדגשות שלי, מ"ר).

כדי לממש את התפישה הזו הוחלט כאמור על הקמת הראמ"ה (שם: 74), ועל שילוב איש הערכה בכל בית ספר (שם: 54). ואכן, אחת התקנות שתיקן משרד החינוך בעקבות הפעילות של כוח המשימה הלאומי (ועדת דוברת, בשנת 2004) הייתה שדרוג של גף הערכה ל"רשות ארצית למדידה והערכה" (ראמ"ה) - גוף סמך אוטונומי האחראי הן להערכה במערכת החינוך והן להכשרת המעריכים.

ההחלטה לשלב איש הערכה בכל בית ספר, וכתוצאה מכך הצורך להכשיר אלפי מעריכים בעבור מערכת החינוך, יצרו מערבולת של הכשרות וניירות עמדה, חלקם ראויים ורובם אינם ראויים (לוי-רוזליס, לפידות ודובר, 2006; לוי-רוזליס ולפידות, 2008).

משרד החינוך והוועדה למחקר יישומי בחינוך שליד המועצה הלאומית הישראלית למדעים, הזדרזו להגיב להחלטות של כוח המשימה הלאומי בהקשר של הערכה. כל אחד מהם פרסם מסמך משלו שמבטא השקפת עולם ומתווה דרכי פעולה. משהוקמה הראמ"ה פרסמה גם היא מסמך משלה. להלן שלושת המסמכים המצויים במרכז השיח על ההערכה בחינוך.

### משרד החינוך

משרד החינוך אימץ את המלצות כוח המשימה הלאומי במלואן, אבל סוגיית המעריך הבית-ספרי עוררה שאלות. מנכ"לית משרד החינוך מינתה ועדת מומחים (ועדת היישום לשילוב הערכה פנימית בבית-הספר), בראשות ד"ר גילי שילד, מנהלת גף מדידה והערכה במשרד, שתבחן את דרכי היישום של ההחלטה לשלב מעריך בכל בית ספר בישראל. הוועדה פרסמה את מסקנותיה במסמך שהוגש למנכ"לית המשרד (ועדת היישום להערכה פנימית, 2004).

ההחלטה הגורפת, לשלב איש הערכה בכל אחד מבתי הספר בישראל, לא התייחסה כלל לסוגיית הכשרתם המקצועית של אלפי המעריכים שיידרשו למערכת. ועדת היישום הדגישה את החשיבות העצומה של הכשרה מקצועית ראויה של איש הערכה הפנימי. המלצה שאומצה מאוחר יותר כאשר הוחלט שרכז הערכה בית-ספרי צריך להיות בעל תואר שני ייעודי בהערכה (חוזר מנכ"ל, 2008). ועדת היישום קיבלה את ההנחות של כוח המשימה הלאומי אבל נתנה להם דגשים משלה (ההדגשות בטקסט שלי. מ.ל.ר.):

**קבלת החלטות** חייבת להתבסס על נתונים אמינים ומשמעותיים שניתן לסמוך עליהם, ומטרת ההערכה היא לספק נתונים/מידע כאלה -- כלומר עליהם לעמוד בקריטריונים מקצועיים מחד, ולשקף את **המטרות הפדגוגיות** של בית-הספר מאידך. לפיכך, הערכה טובה מקיימת את התנאים:

- משרתת את **מקבלי ההחלטות** ועונה לשאלותיהם.
- נעשית בכלים מהימנים ותקפים שניתן לסמוך על המידע שהם מספקים.
- הממצאים ברורים ומובנים, מהווים כלי שימושי ומאפשרים הסקת מסקנות אופרטיביות.
- ניתנת בזמן אמת.

לשון אחר: **כל הערכה חייבת לעמוד בקריטריונים מקצועיים** (ועדת היישום, 2004: 1).

הסוגיה של מקצועיות מודגשת לכל אורכו של המסמך בן עשרת העמודים. המילה *מקצועי* על נגזרותיה נזכרת בו חמישים פעם בהקשר של עבודת רכז ההערכה. הוועדה גם מגדירה מהי המקצועיות הנדרשת, הן מבחינת תכניה והן מבחינה פורמלית. מבחינת התכנים הגדירה הוועדה את הנושאים הבאים: "כדי שלהערכה תהיה תרומה משמעותית חייב המעריך להיות איש מקצוע השולט בידע ובשיטות הקיימים במקצוע, לרבות המיומנויות הבאות (ראו פירוט נוסף בנספח ב):

- "יכולת לשתף את צוות בית-הספר ו/או בעלי תפקידים ולקיים דיאלוג בכל שלבי ההערכה
- בחירת תהליך או שיטת ההערכה המתאימה
- הכרת כלי הערכה ובחירתם לאור בדיקת איכותם
- בחירת האינפורמנטים הנכונים
- איסוף נתונים בדרך המתאימה
- ניתוח ועיבוד נכון של הנתונים
- מתן פירוש נכון לממצאים וכתובת דו"ח בהיר ושימושי

כמו-כן, יש להדגיש, כי הערכה מועילה מחייבת הבנה פדגוגית ודידקטית יסודית על-מנת לתמוך בשיפור תהליכי הוראה ולמידה" (שם: 1).

מבחינה פורמלית הגדירה הוועדה את ההכשרה המקצועית הנדרשת במילים אלה: "רכז הערכה מצוי בצומת מידע חשוב, ובתוקף תפקידו והכשרתו מומלץ שיהיה חלק מצוות הניהול של בית-הספר, ולמילוי תפקידו יוקצו לו שעות... השאיפה לעתיד היא לפתח את מקצוע ההערכה במערכת החינוך, בדומה למודל המקצועי של "יועץ חינוכי" - הן מבחינה מקצועית והן מבחינה ארגונית. כמורים-יועצים משמשים רק אנשים שהוסמכו לכך על-ידי מוסד להשכלה גבוהה והמנחה המקצועי שלהם הוא שפ"י" (שם: 3). ובהמשך: "הוועדה ממליצה על הקמת מסלול אקדמי להכשרת מעריכים שיהיו בעלי תואר שני עם התמחות בהערכה" (שם: 6). כאמור, המלצה זו התקבלה.

הוועדה הייתה מודעת לקושי להכשיר בבת אחת אלפי רכזי הערכה בעלי תואר שני, והציעה שלב ביניים - הכשרה בת 112 שעות לתפקיד "רכז מידע ומשוב". גם המיומנויות הנדרשות מרכז מידע הן רבות ומורכבות, ועיקרן יכולת גבוהה להשתמש במידע ממאגרי נתונים בית-ספריים כדוגמת המנב"ס (מאגר נתונים בית-ספרי) וממאגרי מידע ארציים שונים, הפעלה עקבית ושיטתית של כלי הערכה פנימיים, מקצועיים

ובדוקים הקיימים כבר במערכת החינוך, ועיבוד המידע ממקורות פנימיים וחיצוניים כך שיוכל לקדם את בית הספר בפיתוח של תרבות הערכה ובתהליכים של קבלת החלטות מבוססות נתונים (שם: 4).

ועדת היישום הבחינה בין הערכה פנימית וחיצונית (שם: 2-3), הדגישה את חשיבות המערך המקצועי התומך לאנשי ההערכה הבית-ספריים (עמ' 7), וציינה את כלי ההערכה המקצועיים העומדים היום במערכת לרשות המעריכים. הוועדה הדגישה את רמתם המקצועית הגבוהה ואת הכשרתם הייעודית של אנשי ההערכה הבית-ספריים, קישרה את עבודתם לבית הספר ולצרכיו, הדגישה את החשיבות של ידע פדגוגי ושל בית הספר כצרכן המרכזי של רכז ההערכה. בהתייחס להחלטות כוח המשימה הלאומי, ועדת היישום אימצה את ההערכה ככלי ניהול, פנימי לשם קידומו של בית הספר ולא ככלי פיקוח מערכתי, כפי שדרשו החלטות כוח המשימה הלאומי.

### האקדמיה הישראלית הלאומית למדעים

היוזמה למחקר יישומי בחינוך שליד האקדמיה הישראלית הלאומית למדעים ובשיתוף יד הנדיב, השתתפה אף היא בשיח על ההערכה וביצעה כמה פעולות משמעותיות. בשלב ראשון ערכו שני ימי עיון רחבי היקף (בדצמבר 2004 ובינואר 2005) בהנחייתם של אנשי מדידה חינוכית מחו"ל בעלי שם בינלאומי. בעקבות ימי העיון מונתה ועדת מומחים למדידה ולהערכה בחינוך, שדנה בשאלה מה צריך לדעת איש מדידה והערכה בחינוך. הוועדה הוציאה מסמך "הצעה למסגרת לתכניות לימודים ולפיתוח מקצועי" שהוגש למועצה להשכלה גבוהה (מל"ג), לוועדה לתקצוב ותכנון (ות"ת), למשרד החינוך, ללשכת המדען הראשי, לרשות הארצית למדידה והערכה (הראמ"ה) ולמוסדות אקדמיים (שם). בקיץ 2006 קיימה היוזמה קורס קיץ מרכז בנושא "Evaluating Effectiveness in Education" בראשות פרופסור ברוש מאוניברסיטת פנסילבניה. הקורס נועד למלמדים את הנושא במוסדות אקדמיים, לאנשי הראמ"ה ומשרד החינוך ולמעריכים ממוסדות שונים.

המסמך הנ"ל של הוועדה למדידה ולהערכה בחינוך הוא מקיף ומפורט ומתייחס לאוכלוסיות השונות במערכת החינוך ולסוגי הידע המוצהר והמעשי שעליהן לדעת. אחת מהנחות היסוד של מחבריו היא ש"מטרת העל של מדידה והערכה בחינוך היא לספק לאנשים ולמערכות שונות מידע העשוי לסייע להם ללמוד ולהיות אפקטיביים יותר..." (הוועדה למדידה ולהערכה בחינוך, 2005: 5). בקרה ואחריותיות אינן מוזכרות במסמך, ואת הנחת היסוד השלישית שלו אפשר להבין גם כהתנגדות עדינה לתפישה הזו: "הנחה שלישית קובעת כי במקום להתייחס למדידה ולהערכה באופן מצומצם, כאמצעי להעלאת נכונותם של מורים ומנהלים להיות אחראים להישגי התלמיד, יש להדגיש את האחריות המשותפת המוטלת על כל השותפים במערכת החינוך - לשאוף לתוצאות בעלות ערך אשר כוללות הישגים לימודיים אולם אינן מוגבלות להם בלבד" (שם: 5). חברי הוועדה מזהירים גם מפני התוקף ההשלכתי ומשימוש לא מבוקר בהערכה: "לא רק שמובן שעל משרד החינוך להתוות את המטרות והיעדים למערכת החינוך, אלא שיש להיזהר ממצב שבו בהעדר מטרות שהוגדרו על ידי האחראים, הופכים המודדים והמעריכים לקובעי מטרות בפועל.

בהקשר זה חשוב להזכיר, שידוע שהיחס שבין מדידה והערכה מחד ובין תהליכים חינוכיים ותוצאות מאידך הוא הדדי (מדידה והערכה לא רק שהן משקפות מטרות ותוצאות אלא אף משפיעות עליהן), וכן ידוע שפרקטיקות המקדמות את השגתם של יעדים מסוימים, עלולות במקביל לערער את השגתם של יעדים אחרים. לכן, שימוש מושכל במדידה ובהערכה חייב להביא בחשבון הן את

התועלת לתלמידים, למורים, לבתי הספר ולמערכת החינוך והן את ההשפעות השליליות וההפסדים העלולים להתלוות להן, ולהתבסס על הבנה מעמיקה על אודות הקשרים בין מטרות חינוכיות שונות ובין שיטות שונות של הערכה" (שם: 5).

במסמך הנדון ("מסגרת לתכנית לימודים") יש הבחנה חשובה בין הערכה חינוכית להערכת תכניות ופרויקטים, ובתוך הערכה חינוכית - בין הערכה לצורך למידה (ה"ל"), המעבירה את הממצאים למוערך, להערכה של הלמידה (הש"ל), שהיא חיצונית ומפקחת. מסמך זה הוא הראשון שמתייחס להבדלים שבין הערכה חינוכית כחלק מעבודת ההוראה ובין מדידה חינוכית ופסיכומטריקה. אפשר שמשום כך הוא גם הראשון שמתייחס להכשרה של מורים בשדה בתחום ההערכה: "ייחודם של המורים במערכת (בהקשר הנוכחי) הוא שמחד ההערכה אינה עיסוקם המרכזי (שלא כמו רכזי הערכה ופסיכומטריקאים), ומאידך יש לה השלכות על היבטים רבים של עבודתם. דווקא משום כך ישנה חשיבות רבה לכך שתוקנה למורים תשתית נאותה בתחום ההערכה, גם אם קשה לעתים לבודד את תפקיד המורה כמערך משאר תפקידיו" (שם: 27). המסמך מציע שלושה קורסי בסיס: קורס מבוא (המקנה מושגי יסוד וידע בגישות להערכה, תפקידיה, שיקולים שונים בביצוע ההערכה, והקשר בינה ובין משתנים חינוכיים שונים); קורס בסיס בסטטיסטיקה ובמידה; וקורס שעניינו הערכה בכיתה (classroom assessment) (שם: 17,27).

#### הרשות הארצית למדידה והערכה (הראמ"ה)

בהמלצה של כוח המשימה הלאומי לשיפור החינוך (התכנית הלאומית לחינוך, 2005) הוקמה הראמ"ה ופעילותה אכן מורגשת. אחד הדברים המשמעותיים שעשתה היה קביעת סטנדרטים לידע המקצועי של המעריכים (הרטף, גנור, רום ושלטון, 2007). חוברת המפרטים שקובעת את הסטנדרטים האלה מציינת כי: "תחום ההערכה הפך להיות תחום מרכזי בתהליך החינוכי. תפקיד ההערכה במובן זה נתפס בזיקה גבוהה לתהליך ההוראה-למידה וכחלק בלתי-נפרד ממנו ומחייב ידע ומיומנויות מקצועיות מתקדמות" (שם: 1). מטרת חוברת המפרטים היא "לעשות סדר" או "ליצור תשתית משותפת אשר תאפשר לכלל מעבירי ההשתלמויות בתחום ההערכה ולכלל המשתלמים לדעת את היקף הידע, תחומי ההתנסות והשימושים בפועל (target domain) הנדרשים לכל אחת מההשתלמויות בתחומי המדידה וההערכה בחינוך" (שם: 1). החוברת מפורטת מאוד ומתבססת על שני המסמכים הקודמים (של ועדת היישום ושל הוועדה למדידה והערכה בחינוך).. היא מבחינה בין אוכלוסיות יעד שונות וצורכיהן, בין סוגי הערכה, בין גישות שונות להערכה, ובין תחומי ליבה ותחומים בדרגת חשיבות שנייה. בשונה משני המסמכים הקודמים, המסמך הזה איננו רק בגדר הצעה, אלא שבתחומים מסוימים יש לו תוקף מחייב. כך למשל, במה שאמור במרכזי הפסג"ה. בשנת 2004 התקיימו בהם 77 השתלמויות בהערכה ובשנת 2006 שמונים ושתיים, ובשנה שלאחר פרסום המפרטים ותגבור הפיקוח על תוכני ההשתלמויות ובעיקר על רמתם - היו בהם רק ארבע השתלמויות (לוי-רוזליס ולפידות, 2007), שכן ההשתלמויות שלא עמדו בקריטריונים לא אושרו.

לסיכום סעיף זה נציין שמסמך למסמך<sup>2</sup> ניכרת התפתחות בתפישה של מורכבות התחום. ההחלטות של כוח המשימה הלאומי היו כלליות, ללא ירידה לפרטים (הקמת הראמ"ה ושילוב מעריך פנימי); ועדת היישום חשבה על ה"איך": מהי הדרך הראויה לבצע את ההחלטה; ובשני המסמכים הבאים חשבו גם על ה"מי", ה"מה" וה"כמה": מי הם קהלי היעד, מה עליהם לדעת וכמה. רמת הפירוט במסמך האחרון (חוברת המפרטים של הראמ"ה) רבה יותר מזו שבקודמיו. יש התפתחות גם בהשקפת העולם, החל בתפישה נוקשה של אחריותיות (accountability) ב"תוכנית הלאומית", עבור לתפישה של למידה מהערכה של כל מרכיבי המערכת, וכלה בלמידה עצמית בחוברת המפרטים של הראמ"ה.

בעניין ההכשרה להערכה של סגלי ההוראה – שני המסמכים האחרונים רואים חשיבות רבה בהכשרת המורים ומגדירים בפירוט את תחומי הדעת שהם צריכים לדעת, הן כחלק אינטגרלי מתפקידם כמורים והן בעידן שבו תהליכי הערכה במערכת החינוך הולכים ומתפתחים.

יש לציין שהשיח על ההערכה לא הסתיים במסמכים הכתובים:

היוזמה למחקר יישומי בחינוך ביצעה פעולות נוספות הקשורות לנושא, כמו ארגון ימי עיון וחיבור מסמך הנחיה להערכה ומדידה בגיל הרך (ברזניץ וימין, 2008). הראמ"ה, במקביל לפעולותיה הרבות בתוקף תפקידה כרשות ארצית למדידה והערכה, פועלת רבות לטפח אוריינות ההערכה במערכת החינוך. היא מקיימת ימי עיון רבים לקבוצות אוכלוסייה שונות, השתלמויות שונות, מביאה אנשי שם מחו"ל, נציגיה משתתפים בכנסים ובימי עיון הנערכים על ידי גופים אחרים, והאתר שבבעלותה מעמיד לרשות הציבור חומרים רבים ומגוונים.

גופים שונים מכשירים סגלי הוראה: המכללות להכשרת עובדי הוראה מכשירות פרחי הוראה; האוניברסיטאות מקנות תעודת הוראה; ישנם גופים שיוזמים השתלמויות למורים ב"שדה" בנושאים שונים, והבולטים בהם הם מרכזי הפסג"ה. על פי נתוני האגף להשתלמויות במינהל להכשרה, להשתלמויות ולהדרכה של עובדי הוראה במשרד החינוך, בשנת 2005 התקיימו שבעים ושבע השתלמויות בהערכה במרכזי פסג"ה שונים, במרכז לפיתוח עובדי הוראה בכירים, במכללת בית-ברל, באוניברסיטה הפתוחה, במכון הטכנולוגי לישראל, במכון מופ"ת, בסמינר אורנים ובהשתלמויות בית-ספריות. משך ההשתלמות הממוצעת - 30 שעות.

**להלן ממצאי מחקר**, שנערך בשנים 2006-2008 במסגרת רשת ההערכה של מכון מופ"ת, על הפעילויות הראשונות שנעשו בקשר להכשרה להערכה של סגלי הוראה (לוי-רוזליס, לפידות ודובר, 2006; לוי-רוזליס ולפידות, 2007; לוי-רוזליס ולפידות, 2008).

### מרכזי הפסג"ה

מרכזי הפסג"ה הגיבו בריבוי השתלמויות למורים בפועל. בשנת 2007 הפעילו מרכזי הפסג"ה ברחבי הארץ 84 השתלמויות שונות בהערכה. ממצאי המחקר מלמדים שהמנהלים של מרכזי הפסג"ה תפשו את ההשתלמויות בהערכה כחשובות ונחוצות וגילו מעורבות רבה בהן וציפיות גבוהות מהן.

תכניות הלימודים בהשתלמויות בהערכה היו מגוונות מאוד. מכלל 84 השתלמויות, שמונה נקראו בשם כללי "הערכה" או "הערכה ומדידה", עשרים ותשע אחרות שילבו למידה והערכה ("דרכי הוראה והערכה"

<sup>2</sup> התכנית הלאומית לחינוך (כוח המשימה הלאומי), המלצות ועדת היישום להערכה פנימית (משרד החינוך, במסגרת כוח המשימה הלאומי), הצעה למסגרת לתכניות לימודים ולפיתוח מקצועי (היוזמה למחקר יישומי בחינוך), וחוברת המפרטים (הראמ"ה)

הוא השם הנפוץ), ארבע יועדו לרכזי/נאמני הערכה, והשאר (43 השתלמויות) דנו בנושאים שונים ומגוונים: "הערכה בגן הילדים", "הערכה בלימוד מדעי כדור הארץ", "תרבות הערכה", "אקלים בית ספרי בדגש על הערכה" ועוד.

המרצים המלמדים בהשתלמויות הם בעלי רקע דיסציפלינרי מגוון. רובם ככולם עוסקים בהערכה ומכירים בחשיבותה. הם מבקשים לראות במשתלמים סוכני שינוי, סוכני הטמעה של תרבות ההערכה הבית-ספרית. יוזמי ההשתלמויות מגוונים אף הם. 39 השתלמויות מתוך ה-84 במרכזי הפסגה היו ביוזמת מנהליהם. את כל השאר יזמו מפמ"רים, יחידות מטה שונות במשרד החינוך, וגופים אחרים מחוצה לו כמו הטכניון, האוניברסיטה הפתוחה, מט"ח, אוניברסיטת בר אילן, אוניברסיטת תל אביב, מכון וייצמן ומכללת אחוה.

מתוך הממצאים נראה שתרומת ההשתלמויות הייתה חלקית, וזאת מכמה סיבות - חלקן טכניות וחלקן מהותיות. הבעיות הטכניות נגרמו בעיקר בגלל מספר השעות הקטן שהוקדש לנושא כה רחב ומורכב המיועד לציבור משתלמים חסר רקע בתחום. מיעוט השעות גרם לרדידות של ההכשרה בתחום - השתלמויות קצרות שכיסו מעט, ונבנו על פי תפישתו של המרצה. במקביל, ומכיוון שההערכה מחוברת כמעט לכל תחום בעשייה החינוכית, גופים רבים במערכת יוזמים השתלמויות למורים בתחום ספציפי ממוקד של הערכה. יש לכך יתרונות רבים, אבל במבט כללי נוצר פיצול וקטוע של התחום ובלבול גדול, שכן לא תמיד ברור על מה מדברים, לאיזה תחום שייך הידע המסוים, לאיזו גישה, לאיזו השקפת עולם.

בהיעדר מדיניות ברורה מה לומדים ולצורך מה, כל מערך ההכשרה נראה די כאוטי (לוי-רוזליס ולפידות, 2007). את מקום המדיניות החסרה תפסה חוברת המפרטים של הראמ"ה, שמנסה לתת תשובות מסודרות לכל השאלות האלה, ויש לה גם כוח מינהלי לקבוע נורמות וכללים להערכה ראוייה.

### **מדיניות ההכשרה להערכה במכללות להכשרת עובדי הוראה**

ההכשרה להערכה של פרחי הוראה במכללות להכשרת עובדי הוראה ובאוניברסיטאות, נבדקה ב-2006 וב-2008 (לוי-רוזליס ולפידות, 2007, 2008). הממצאים רבים ומגוונים (ראו פירוט בדו"חות המלאים), אבל מדיניות ברורה לא מצאנו. הכשרה זו שונה ממכללה למכללה, מתפרסת על פני סוגים שונים של תכניות לימוד ופונה לקהל מגוון. באופן כללי נחלקות התכניות לשתי קבוצות: האחת - תכניות המיועדות לפרחי הוראה, קורסים ייעודיים להערכה, קורסים תלויי דיסציפלינה, במסגרת לימודי פדגוגיה וחינוך ובמסגרת סמינריונים דידקטיים ומחקריים. השנייה - תכניות הנכללות במסגרת לימודי ההמשך: השתלמויות למורים ומנהלים בפועל ובמסלולים לתואר שני.

היקף הקורסים מצומצם, רובם סימסטריאליים ואינם מלווים בתרגול. עם זאת, הם עשירים בתכנים: הערכה מעצבת ומסכמת, הל"ה (הוראה-למידה-הערכה), הערכת הישגים, בניית כלים, בניית מחוון, הערכה חלופית, כלי הערכה חלופיים (מטלת ביצוע, תלקיט, עבודת חקר), תרבות הערכה בית-ספרית, ניתוח נתונים, דיווח תוצאות, דילמות אתיות, הערכה פנימית וחינוכית וכדומה. בחלק מהקורסים נמצאו כל הנושאים האלה אך לכל אחד מהם הוקדש שיעור אחד.

המרצים שמלמדים הערכה באים מדיסציפלינות שונות ולרובם אין הכשרה בהערכה או במחקר. כולם חושבים שהנושא חשוב, אך הידע של רבים מהם באשר למהות התחום חלקי ומבלבל.

במכללות שבדקנו לא מצאנו חשיבה על מדיניות, על גוף ידע ברור, על סדר עדיפויות לגבי מה חשוב ללמד ומה לא, מתי וכמה. במכון מופ"ת פעלה קבוצת חשיבה וולונטרית של נציגי מכללות שונות, אבל עדיין אין לה השפעה בשטח.

לתוכנית הלאומית לחינוך (דו"ח דוברת) הייתה השפעה מסוימת על כמות הקורסים המיועדים לפרחי הוראה. בשנה הראשונה של המחקר (2006) מצאנו בחמש עשרה מכללות כ-20 קורסים ייעודיים וכ-30 קורסים להערכה תלויי דיסציפלינה. בבדיקה חוזרת כעבור שנתיים (2008) מצאנו ב-18 מכללות 28 קורסים ייעודיים להערכה ו-40 קורסים תלויי דיסציפלינה. מספר הקורסים בלימודים לתואר שני גדל באותה עת כמעט פי שניים - מתשעה קורסים לשבעה עשר. לא נמצאו הבדלים בתכנים או בהיקף השעות.

### מדיניות ההכשרה להערכה באוניברסיטאות

בכל התכניות שמקנות תעודת הוראה יש קורסים בהערכה. נבחנו שמונה קורסים כאלה, בחמש אוניברסיטאות ובטכניון, רובם קורסי חובה. המצב באוניברסיטאות דומה לזה שבמכללות, למעט העובדה שהמרצים המלמדים הערכה באים מהתחום, והתפישה שלהם את נושא ההערכה ואת הרציונל של מבנה הקורס היא קוהרנטית ומדויקת יותר. אבל גם כאן, ההחלטה על מבנה הקורס ועל תכניו היא בידי מלמד הקורס.

### סיכום

הערכה איננה "דיסציפלינה". היא קיימת בכל תחום, החל בחקלאות ובבריאות, דרך מיזמים הנדסיים, פעולות חברתיות וקהילתיות, וכלה בחינוך, ולכן אין לה גוף ידע קנוני ברור. דווקא משום כך חשוב לקבוע מדיניות ולהחליט מה חשוב לעובדי הוראה ומה אינו חשוב. הוסף לכך את תפקידיה השונים של הערכה בכלל ושל הערכה בחינוך בפרט, ואת הגישות השונות הקיימות להערכה, ותבין את הבלבול השורר בתחום. העדר מדיניות ברורה באשר למקומה של ההערכה בתפקידים ובהכשרתם של עובדי הוראה, גורר ארבע בעיות מרכזיות:

1. אין החלטה או חשיבה על גוף הידע החשוב למורים, לפיכך גוף הידע לא ברור ויש בלבול וחוסר הבחנה בין גישות שונות ותפקידים שונים של הערכה.
2. קיים מחסור בכוח אדם מיומן בתחום זה.
3. הזמן המוקדש להכשרה להערכה מצומצם מדי.
4. אין תכנית הכשרה מובנית.

דו"ח דוברת תרם להעלאת נושא ההערכה למרכז הבמה החינוכית ועורר את השיח הציבורי על נחיצותה של ההערכה במערכת החינוך. כאמור במבוא, נושא ההערכה היה קיים במערכת קודם לכן, אך היה ברובו ספורדי, תלוי אגף או אדם. הדו"ח והמלצותיו זימנו דיון רחב בנושא ההערכה בחינוך. נקודת המוצא של כוח המשימה הלאומי (ועדת דוברת) הייתה תפישה של אחריותיות (accountability) ופיקוח. תפישה קיצונית זו, כפי שבוטאה בתכנית הלאומית לחינוך, 2005, דעכה מעט, אבל סוגיית האחריותיות עדיין מלווה את השיח על ההערכה. מנגד ישנה תפישה שמדברת על הערכה כאמצעי



ללמידה. גם הרשות הארצית למדידה והערכה (הראמ"ה) אימצה כמוטו את הסיסמה "מדידה לצורך למידה".

מהתבוננות במסמכים הרבים שצוטטו לעיל, מניתוחם ומניתוח הנעשה בתחום של הכשרת עובדי הוראה, מתברר שאין חשיבה מערכתית ואין מדיניות באשר לידע המקצועי הנדרש למורים בנושא הערכה. לפני פרסום התכנית הלאומית לחינוך לא תפסה ההערכה מקום של ממש בהכשרה של עובדי הוראה כחלק מתפקידו המקצועי של מורה, לפחות בהערכת הישגי תלמידיו, ולא העסיקה את השיח הציבורי. פרסום התכנית הלאומית תיקן את המעוות, אלא שהשיח הציבורי התמקד בסוגיית ההערכה בחינוך ורק מעט מאוד בשאלה מה צריכים המורים לדעת. מהמסמכים השונים שניתחתי יש רק שניים שדנים בידע הנדרש לאנשי חינוך ולמורים: המסמך "הצעה למסגרת תכניות לימודים ולפיתוח מקצועי - מדידה והערכה בחינוך" של היוזמה למחקר יישומי בחינוך (הוועדה למדידה והערכה בחינוך, 2005), ו"מדריך לתכנון השתלמויות במדידה והערכה בחינוך לשם פיתוח מקצועי של עובדי הוראה" מטעם הראמ"ה (הרטף ואחרים, 2007). המסמך הראשון הוא בגדר המלצה בלבד, ולשני יש השפעה מוגבלת, בעיקר על איכות השתלמויות של המורים בפועל אך לא על עצם קיומן או על כמותן. עוד פחות מכך משפיעים המסמכים על פעולות מערכתיות, אקטיביות ורחבות היקף להכשרת עובדי הוראה.

במכללות האקדמאיות להכשרת עובדי הוראה, שהן המכשירות העיקריות של מורים בישראל, לא נוסחה (נכון למועד כתיבת המאמר, 2010) שום מדיניות להכשרת עובדי הוראה להערכה, ומה שנעשה בהן הוא מקומי, חלקי ותלוי במי שנמצא בשטח.

על אף שידע מקצועי בהערכה חיוני למורים, בהיותו חלק מהגדרת הידע המקצועי שלהם ומזהותם המקצועית, אין היום במדינת ישראל שום מדיניות בנושא זה.

## ביבליוגרפיה

- אורמיאן, חיים (עורך) (1979). עמנואל יפה – במעגלי החינוך. ירושלים: משרד החינוך והתרבות.
- בירנבוים, מנוחה (1997). חלופות בהערכת הישגים. תל אביב: רמות.
- בירנבוים, מנוחה (2006). הערכה אלטרנטיבית – לא כל הנוצץ זהב הוא! חינוך בראש העל-יסודי, 1, 23-19.
- ברזניץ, צביה וימין, רעות (עורכות) (2008). אבחון, מדידה והערכה, גן-ב', תמונת מצב והמלצות. **היזמה למחקר יישומי בחינוך**. ירושלים: האקדמיה הלאומית הישראלית למדעים, משרד החינוך והתרבות והספורט, יד הנדיב.
- בריטברד, רוני וטביביאן-מזרחי, מיכל (2003). שינויים שחלו במבנה מערכת החינוך בישראל והצעות לשינויים שעלו במהלך השנים. מסמך רקע, מוגש ליו"ר ועדת החינוך והתרבות של הכנסת. ירושלים: הכנסת, מרכז מחקר ומידע.
- האקדמיה הישראלית הלאומית למדעים (2005). מה צריכים לדעת העוסקים בהערכה בחינוך: סמינר. היוזמה למחקר יישומי בחינוך, כפי שאוחזר ב-24 באפריל 2010: <http://www.academy.ac.il/data/projects/34/E13.pdf>
- הוועדה למדידה ולהערכה בחינוך (2005). הצעה למסגרת לתכניות לימודים ולפיתוח מקצועי - מדידה והערכה בחינוך. היזמה למחקר יישומי בחינוך.. ירושלים: האקדמיה הלאומית הישראלית למדעים, משרד החינוך והספורט, יד הנדיב.
- הלוי, אסתר (2009) התהוות שדה הערכה בישראל – ההיבט המוסדי. חיבור לקבלת תואר "מוסמך למדעי הרוח והחברה" (M.A). באר-שבע: אוניברסיטת בן-גוריון בנגב.
- ועדת היישום לשילוב הערכה פנימית בבית הספר (2004). המלצות: במסגרת יישום המלצות כוח המשימה הלאומי לקידום החינוך בישראל. ירושלים: משרד החינוך והספורט.
- הרטף, חגית, גנור, מרגלית, רום, אמירה, שלטון, חני (2007). מדריך לתכנון השתלמויות במדידה והערכה בחינוך לשם פיתוח מקצועי של עובדי הוראה. הרשות הארצית למדידה ולהערכה (הראמ"ה). משרד החינוך והתרבות.
- התכנית הלאומית לחינוך (2005). כוח המשימה הלאומי לקידום החינוך בישראל (דוברת, שלמה, יו"ר). ירושלים: משרד החינוך.
- ואן גלדר, אלעד (2003). מסמך רקע לדיון בנושא הכשרת מורים בישראל: מבנה תכניות הלימוד במוסדות להכשרת מורים. מוגש לוועדת החינוך והתרבות. ירושלים: הכנסת.
- ועדת אריאב (2006). החלטת המועצה להשכלה גבוהה מיום 21.11.2006 בנושא "מתווים מנחים להכשרה להוראה במוסדות להשכלה גבוהה בישראל" – דו"ח "ועדת אריאב". ירושלים: המועצה להשכלה גבוהה.

ועדת החינוך של הכנסת (2002). *פרוטוקול מס' 475*. ירושלים: הכנסת.

ועדת הקבע למסלולים אקדמיים במוסדות להכשרת עובדי הוראה (1981), *דגם מנחה לתוכנית לימודים לתואר 'בוגר בהוראה' (B.Ed.)*, 13 בדצמבר 1981; ועדת הקבע למסלולים אקדמיים במוסדות להכשרת עובדי הוראה. ירושלים: המועצה להשכלה גבוהה.

חוזר מנכ"ל (2000). *נוהל ההערכה של המתמחה בהוראה (סטאז'ר)*. סעיף 2-8.2 (א). אפריל 2000 ירושלים: משרד החינוך.

חוזר מנכ"ל (2008). *רכז הערכה בית ספרי*. חוזר מנכ"ל, 10, סעיף 1-1.8. ירושלים: משרד החינוך.

יוגב, אברהם, ליבנה, עידית, ופניגר, יריב (2009). סקר בינלאומי של הישגים חינוכיים והגלובליזציה של מטרות החינוך. *מגמות*, 46 (3).

לוי, אריה (ללא תאריך). *הלכה ומעשה של תכניות לימודים. הלכה למעשה - אוסף מאמרים*. ירושלים: משרד החינוך, המזכירות הפדגוגית, האגף לתכנון תכניות לימודים. אוחר ב-24.4.10.

[http://www.education.gov.il/tochniyot/Limudim/halacha/kategory\\_5.htm](http://www.education.gov.il/tochniyot/Limudim/halacha/kategory_5.htm)

לוי-רוזליס, מירי (2005). קיברנטיקה – מערכת עבודה משותפת בין מעריכים פנימיים וחיצוניים. בתוך: זהבה' בן-עמי (עורכת), *מדידה והערכה, אחריותיות ושקיפות ככלי ניהול לשירות מערכת החינוך*. תל אביב: מכון מופ"ת.

לוי-רוזליס, מירי (2008). *איך מרימים את ההר או אנטומיה של כישלון*. הרצאה בכנס העשור של איל"ת – אגודה ישראלית להערכת תכניות, פברואר 2008, אוניברסיטת בר-אילן.

לוי-רוזליס, מירי, לפידות אורית' ודובר, מ' (2006). *ההערכה כאמצעי בידי המורה /או בידי המערכת, והשפעותיה האפשריות על ההגדרה, על ההכשרה ועל תפקוד המורה הראוי*. דו"ח מחקר, חלק א. רשות המחקר, במסגרת רשת עמיתית מחקר. תל אביב: מכון מופ"ת.

לוי-רוזליס, מירי ולפידות, אורית (2007) *ההכשרה להוראה של מורים מכהנים במרכזי הפסגה* (תשס"ז). דו"ח מחקר חלק ב, במסגרת רשת עמיתית מחקר. תל-אביב: מכון מופ"ת

לוי-רוזליס, מירי ולפידות, אורית (2009) *הכשרה להערכה של עובדי הוראה בישראל*. דו"ח מחקר מסכם לשלוש שנות מחקר, במסגרת רשת עמיתית מחקר. תל-אביב: מכון מופ"ת

לוי-רוזליס, מירי ולפידות, אורית (2010). *שדה ההערכה בישראל וזהותם המקצועית של המורים – היש למכללות להכשרת עובדי הוראה סיכוי במשחק? דפים*, 49 עמ' 147-119 תל אביב: מופ"ת.

מירו, אירית (2009). *המשוב הבית-ספרי כמנוף לשינויים בית-ספריים ומערכתיים*. חיבור לשם קבלת תואר דוקטור לפילוסופיה. באר-שבע: אוניברסיטת בן-גוריון בנגב.

נידרלנד, דורון, דרור יובל, והופמן עמוס (2009) *דמות המורה בראי הכשרת המורים: מבט היסטורי*. דו"ח מחקר, רשת עמיתית מחקר. תל אביב: מכון מופ"ת

נידרלנד, דורון והופמן, עמוס' (2010). דמות המורה בראי הכשרת המורים: מבט היסטורי. דפים, 49 עמ' 86-43 תל אביב: מופ"ת.

פרידמן, יצחק (2005). *מדידה והערכה של תכניות חברתיות וחינוכיות*. ירושלים: מכון הנרייטה סאלד.

**Alkin, M., (1970). Products for improving educational evaluation. *Evaluation Comment*, vol. 2 (3), 1-15.**

Berger, P. L.&. Luckmann, T (1966). *The social construction of reality: A treatise in the sociology of knowledge*. Anchor Books: Garden City, NY.

Bloom, B. S. et al, (eds.) (1971). *Handbook of Formative and Summative Evaluation of Student Learning*. McGraw Hill, New-York 1971a.

Bray, M. & Mukundan, M.V. (2003). *Management and governance for EFA: Is decentralisation really the answer? Paper commissioned for the EFA Global Monitoring Report 2003/4, The Leap to Equality*". Hong Kong: Comparative Education Research Centre, Faculty of Education, the University of Hong Kong.

Chianca, T. (2004). The global evaluation community, the Brazilian community, the Brazilian Evaluation Network and the WMU Evaluation Center. Presentation at the Evaluation Café Presentation The Evaluation Center, Western Michigan University.

Cochran-Smith, m. (2003). Learning and unlearning: The education of teacher educators. *Teaching and Teacher Education*, 19, 5-28

Cousins, J.B., & Aubry, T., (2006) Roles for government in evaluation quality assurance: Discussion Paper. Ottawa: Centre of Excellence for Evaluation, Treasury Board of Canada, Secretariat.

Covaleskie, J.F (1994). The educational system and resistance to reform: The limits of policy. *Education Policy Analysis Archives*, 2 (4). Retrieved on August, 28<sup>th</sup> 2011 from: <http://epaa.asu.edu/epaa/v2n4.html>

Cronbach, L.J., (1983). *Designing evaluation of educational and social programs*. San Francisco: Jossey Bass.

Cuban, L. (2007). Hugging the middle: Teaching in an era of testing and accountability. *Education Policy Analysis Archives*, 15 (1). Retrieved on **August 28<sup>th</sup> 2011** from <http://epaa.asu.edu/epaa/v15n1/>

Datta, L.-E. (2003). The evaluation profession and the government. In: T. Kellaghan & D. L. Stufflebeam (Eds.), *International Handbook of Educational Evaluation*, 345-360. Boston: Kluwer.

Darling-Hammond, L. & Snyder, J. (1992). Reframing accountability: Creating learner-centered schools. In: A. Lieberman (Ed.), *The changing contexts of teaching (Ninety-first Yearbook of the National Society for the Study of Education)*, Chapter 2, pp. 11-36. Chicago, IL: University of Chicago Press.

Darling-Hammond, L. (2000). How teacher education matter. *Journal of Teacher Education*, 51 (3), 160-173

De Jager, P. (2001). Resistance to change: A new view of an old problem. *The Futurist*, 35 (3), 24-27.

- Doran, H. C., (2003). Evaluating the consequential aspect of validity on the Arizona instrument To measure standards. Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, April 21-25, 2003).
- Elliott, J., (1999). Global and local dimensions of reform in teacher education (Editorial Introduction). *Teaching and Teacher Education* 15 (2), 133-141.
- Ertle, H. (2006). Educational standards and the changing discourse on education: The reception and consequences of the PISA study in Germany. *Oxford Review of Education*, 32 (5), 619–634.
- Feiman-Nemser, S & Remillard, J., (1996). Perspectives on learnings to teach. In: F.B. Murray (Ed.) *The Teacher Educators' Handbook*, 63-91. San Francisco: CA.
- Fullan, M. G. (1993). Why Teachers Must Become Change Agents? *Educational Leadership* 50(6). As retrieved on August 24<sup>th</sup>, 2011 from: [http://www.michaelfullan.ca/Articles\\_98-99/03\\_93.htm](http://www.michaelfullan.ca/Articles_98-99/03_93.htm)
- Fullan, M. G. (2001) *Leading in a Culture of Change*. San Fransisco: Jossey-Bass**
- Goodlad, J., (1995). *Teachers for our Nation's Schools*. San Francisco: Jossey-Bass.
- Green, T.F. (1994). Policy questions: A conceptual study. *Education Policy Analysis Archives*, 2 (7). Retrieved on August 28<sup>th</sup> 2011 from: <http://epaa.asu.edu/epaa/v2n7.html>
- Guba, E.G. & Lincoln, Y.S... (1989). *Fourth Generation Evaluation*. . London: Sage Publications.
- Greenwald, A. G., Nosek, B. A., and Sriram, N. (2006). Consequential validity of the implicit association test: comment on Blanton and Jaccard. *American Psychology*, 61 (1), 27-41
- Gussman, T. K. (2005, May). Improving the professionalism of evaluation. Ottawa: Centre for Excellence in Evaluation, Treasury Board Secretariat of Canada.
- Hall, P.M. & McGinty, P.J. W. (1997). Policy as the transformation of intentions: Producing program from statute. *The Sociological Quarterly*, 38 (3)' 439-467.
- Hamilton, L. S., Stecher, B.M., and Klein S. P. (Eds.)(2002). *Making Sense of Test-Based Accountability*. Santa Monica, CA: Rand Corporation.
- Heim, M. (February, 1995). "Accountability in Education." (unpublished paper submitted to the Hawai's Department of Education in partial fulfillment of professional improvement leave requirements). Honolulu, HI: Hawai'i State Department of Education, Office of the Superintendent.
- Henry, M., Lingard, B., Rizvi, F., & Taylor, S. (2001). *The OECD, globalisation and education policy*. Oxford: Pergamon.
- Henry, G.T., & Mark, M. M., (2003). Beyond use: Understanding evaluation's influence on attitudes and actions. *The American Journal of Evaluation*, 24 (3), 293-314.
- Howe, K. R. (2005) The question of educational science: Experimentism vs. experimentalism *Educational Theory*, 55 (3), 307–321.
- Jennings, J. & Rentner, D. (2006). Ten big effects of the No Child Left Behind Act on public schools. *Phi Delta Kappan*. October, 2006. Retrieved February 05, 2009, from [http://www.pdkintl.org/kappan/k\\_v88/k0610jen.htm](http://www.pdkintl.org/kappan/k_v88/k0610jen.htm)

- Jones, L. V. & Olkin, I.(Eds.) (2004). The nation's report card: Evolution and perspectives. Bloomington: *Phi Delta Kappa Educational Foundation*.
- Kellaghan,T,& Greaney, V, (2001) *Using assessment to improve the quality of education*. Paris:UNESCO.
- Kohn, A., (2000) The case against standardized testing: Raising the scores, ruining the schools. Portsmouth: Heinemann.
- Koretz, D. & Barton, K. (2004). Assessing students with disabilities: Issues and evidence. *Educational Assessment* 9 (1-2), 29–60.
- Koretz, D., and Barton, K. (2003-2004). Assessing students with disabilities: Issues and evidence. *Educational Assessment*, 9(1&2), 29-60. (2004)
- Lazear, E.P. (2005). *Speeding, tax fraud, and teaching to the test*. CSE Report, 659. Los Angeles: Center for the Study of Evaluation, National Center for Research on Evaluation, Standards, and Student Testing, Graduate School of Education & Information Studies, University Of California. **7.7.11**
- Levin–Rozalis, M. (2003). The differences between evaluation and research. *The Canadian Journal of Program Evaluation*, 18 (2), 1-31.
- Levin-Rozalis, M. and Lapidot, O., (under review). Evaluation in teacher training colleges in Israel: Do teachers stand a chance? *Journal of Assessment and Accountability Systems in Educator Preparation*.
- Levin-Rozalis, M.. and Rosenstein B. (2005). The role of the evaluator in the process of an organizational learning. *The Canadian Journal of Program Evaluation* 20 (1), 81-104.
- Levin-Rozalis, M. and Shochot-Reich, E. (2008). Professional identity of evaluators in Israel. *The Canadian Journal of Program Evaluation*. 141-178
- Levin-Rozalis, M., Rosenstein, B. and Cousins, B. J. (2009). Precarious balance: Educational evaluation capacity building in the era of globalization (cp. 11). In: Ryan, K & Cousins, J. B., *Sage International Handbook on Educational Evaluation*. Thousand Oaks: Sage
- Levin-Rozalis, M. and Lapidot, O. (2010). Evaluation in teacher training colleges in Israel: Do teachers' colleges stand a chance? *Journal of Assessment and Accountability Systems in Educator Preparation*.
- Love, A.J., Russon, C. (2000). Building a worldwide evaluation community: Past, present, and future. *Evaluation and program planning*, 23, 449-459.
- Love, A. J. (1991). *Internal evaluation: Building organizations from within*. Newbury Park, CA: Sage Publications Inc.
- Love, A. (2005). Internal Evaluation. In: S. Mathison (Ed.), *Encyclopedia of Evaluation*. Thousand Oaks, CA: Sage.
- McDonnell, M. L., (2005). No Child Left Behind and the federal role in education: Evolution or revolution? *Peabody Journal of Education*, 80 (2), 19-38.
- Messick, S. J. (1989). Validity. In: Linn, Robert (Ed.), *Educational Measurement, 3rd Edition* (13-103). Washington, DC: American Council on Education.
- Messick, D. M., H. Wilke, M. B. Brewer, R. M. Kramer, P. E. Zemke, and L. Lui. 1983. Individual Adaptations and Structural Change as Solutions to Social Dilemma. *Journal of Personality and Social Psychology* 44:294-309.
- Morrow, R.A., & Torres, C.A. (2000). The state, globalization and educational policy. In: N.C. Burbules & C.A. Torres (Eds.), *Globalization and Education: Critical Perspectives* (27-56). New York and London: Routledge.

- Nevo, D. (2001). School evaluation: Internal or external? *Studies in Educational Evaluation*, 27, 95-106.
- Patel, M. and Russon, C. (1998). Appropriateness of the Program Evaluation Standards for Use in Africa. Paper presented to the African Evaluation Association.
- Patton, M. Q. (1997). *Utilization-focused evaluation*, 3<sup>rd</sup> Edition. Thousand Oaks, CA: Sage Publications.
- Richardson, V. (1997). Constructivist teaching and teacher education: Theory and practice. In: V. Richardson (Ed.), *Constructivist teacher education, building a world of new understanding* (3-14). The Flamer Press: Washington: D.C.
- Romainville, M. (2002). *Du bon usage de PISA*. *La Revue Nouvelle*, 115 (3-4), 86-99.
- Sarason, S., (1998). *Political leadership and educational failure*. San Francisco: Jossey Bass.
- Scriven, M. (1967). The methodology of evaluation. in: R.E. Stake (Ed.), *Curriculum Evaluation, AERA Monograph Series on Evaluation, No. 1*. Chicago: Rand McNally.
- Shohami, E. (2001). *The power of tests: A critical perspective on the use of language tests*. London: Pearson Education
- Simola, H. (2005). The Finnish miracle of PISA: historical and sociological remarks on teaching and teacher education. *Comparative Education*, 41 (4), 455-470.
- Smith, M.S., O'Day, J. & Fuhrman, S.H. (1994). State policy and systemic school reform. In: C.M. Reigeluth, & R.J. Garfinkle. (Eds.). *Systemic change in education*. Englewood Cliffs: Educational Technology Publications
- Stake, R. E. (1967). The countenance of educational evaluation. *Teachers College Record*, 68.
- Stake, R. E. (2007). NAEP, Report cards and education: A review essay. *Education Review*, 10 (1). Retrieved February 8, 2007, from <http://edrev.asu.edu/essays/v10n1index.html>.
- Spillane, J. P. (2004). *Standards deviation: How schools misunderstand education policy*. Cambridge, Massachusetts: Harvard University Press.
- Schwartz, R. (1998). The politics of evaluation reconsidered. *Evaluation*, 4 (3), 294-309.
- Toulemonde J. 2000. "Evaluation culture(s) in Europe: differences and convergence between national practices," in *Vierteljahrshefte zur Wirtschaftsforschung* 69: 350-357.
- Tyler. R. W. (1942). General statement on evaluation. *Journal of Educational Research*, 35, 492-501.
- Weiss, C. H. (1991). Evaluation research in the political context: Sixteen years and four administrations later. In: M. W. McLaughlin & D. C. Phillips (Eds.), *Evaluation and education: At a quarter century* (211-231). Chicago: University of Chicago Press.
- Weiss, C. H. (1999). The interface between evaluation and public policy. *Evaluation*, 5 (4): 468-486.
- Wholey, J. S. (1983). *Evaluation and effective public management*. Boston: Little Brown.

Wholey, J. S. (1994). Assessing the feasibility and likely usefulness of evaluation.. In: J. S. Wholey, H. P. Hatry & K. E. Newcomer (Eds.), *Handbook of Practical Program Evaluation* (15–39). San Francisco: Jossey Bass.