

את מי משרתת ההערכה ולאילו צרכים?

מדיניות ההכשרה להערכה של סגלי הוראה

מירי לויין-רוזלים

פרק 5

תקציר

הערכה היא חלק בלתי נפרד ממערכת החינוך בכלל וממעשה ההוראה בפרט. למרות זאת, בכל הועדות שעסקו באופן נרחב בתהליכי הכשרת המורים, ושחשבו וניתחו לעומק מה הם צריכים לדעת, לא מוזכרת ההערכה מכל סוג שהוא. לא הערכת התלמידים על ידי המורים, ולא הערכת המורים את עבודת עצמם. כל אלו היו לגמרי מחוץ לשדה הראייה, כך עד להקמת "כוח המשימה הלאומי לקידום החינוך" ("ועדת דוברת"), שמסקנותיה עוררו שיח ער ואפילו מאבקים סביב סוגיית ההערכה, יחד עם שתי השלכות מעשיות: הקמת הראשות הארצית למדידה ולהערכה, וההחלטה לשילוב רכז הערכה בכל אחד מבתי הספר בישראל. עם זאת, בכל האמור בידע המקצועי של מורים בתחום ההערכה, עולה כי אין חשיבה מערכתית או מדיניות הנוגעת להכשרת עובדי הוראה, וההערכה נשארת מחוץ לתהליכי ההכשרה של המורה בישראל.

מדיניות

במדיניות בכלל ובמדיניות חינוכית בכלל זה, הכוונה היא לפעולה מכוונת או שינוי מתוכנן, שנובעים מתוך תפיסת אי נחת של מצב קיים (Hall & Green, 1994; Covaleskie, 1994; Spillane, 2004; McGinty, 1997).

מקומה של ההערכה ביישום מדיניות

אחד ממנגנוני הפיקוח היעילים על יישום מדיניות, או אף על דחיפתה – זו הערכה. ואכן בשנים האחרונות אנחנו מוצאים הערכה במקומות רבים שבהם יש רצון לשנות דפוסי התנהגות קיימים ולהחדיר דפוסי חדשים. בשביל לחזק את הטענה הזו, די אם נסתכל היום בקורה בתחום ההערכה בעולם. גופים פינאנסיים גדולים וחזקים וגופים אחרים בעלי השפעה עולמית מעורבים היום באופן משמעותי ביותר בשדה ההערכה. למשל, הבנק העולמי שמקים ארגוני הערכה בכל אפריקה, הממשל הפדראלי בארצות הברית שמכתיב סטנדרטים להערכה ולמחקר, האיחוד האירופאי ששם נציבי הערכה במדינות המצטרפות לאיחוד, ממשלות ברוב מדינות אירופה שמפעילות יחידות הערכה, האו"ם וסוכנויות שונות שלו שגם מפעילים הערכה וגם מקיימים הכשרות למעריכים, ה-OECD ששותף לארגוני הערכה בינלאומיים חזקים. כמו גופי הערכה רבי עצמה (דוגמת NONIE התארגנות שמאחדת בתוכה 70 ארגונים ביניהם מחלקת ההערכה של האו"ם, הבנק העולמי, ה-OECD-DAC, יחד עם ארגונים לאומיים במדינות מתפתחות) שמכתיבים במידה רבה את שאלות ההערכה ואופני ההערכה בחלק ניכר מהגופים האלה (ראו למשל: <http://www.worldbank.org/ieg/nonie/> או <http://www.oecd.org/oecdworldforum>).

Chianca, 2004; Cousins & Aubry, 2006; Datta, 2003; Gussman, 2005; Kellaghan, & Greaney, 2001; Levin-Rozalis, Rosenstein and Cousins, 2009; Love & Russon, 2000; (Toulemonde, 1999).

כל אלו לא עושים זאת רק משיקולי מקצוע או רק מכוונות טובות. המטרות שלעיתים הן מוצהרות חלקית ולעיתים אינן מוצהרות כלל, הן מטרות של פיקוח, של הטמעת סטנדרטים שמתאימים או חשובים לגופים המתערבים, של השתתת תרבות שלטונית שהם יכולים לחיות איתה, של השפעה על התנהלות של גופים, ארגונים, מערכות ואף ממשלות, פיקוח על זרימת הכספים ואופני הוצאתם, ובמקרים לא מעטים, גם קולוניאליזם תרבותי.

מערכות החינוך בארץ ובעולם איננה יוצאת מן הכלל במה שאמור בהערכה כמנגנון פיקוח יעיל. תחום זה ובמיוחד מבחנים קובעים משמשים עבור גופים מרכזיים כלי לשליטה ולהכתבה של סדר היום במערכות חינוכיות ובחברה ככלל (מירו, 2009; Henry & Mark, 2007; Cuban, 2007; Jennings & Rentner, 2006; 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 ארגונים על-מדינתיים כמו הארגון לשיתוף פעולה כלכלי ה-OECD, שאחראי על מבחני ה-PISA (Program for International Student Assessment), האגודה הבינלאומית להערכת הישגים לימודיים ה-IEA שאחראית על מבחני ה-TIMSS (The International Mathematics and Science Study), הבנק העולמי והאיחוד האירופאי, יחד עם גופים כמו אונסק"ו הם שדוחפים את המבחנים האלה במטרה ואחרים. מטרתם היא להשפיע על מעצבי המדיניות במדינות השונות, להכין את התלמידים לכלכה הגלובלית (Henry, Lingard, Rizvi and Talor, 2001), בלי שלהם עצמם יש אחריות על הנעשה באותן מדינות (יוגב, לבנה ופניגר, 2009). חוקרים שונים בתחום מתארים את הריבוד הבין מדינתי והשלכותיו השליליות (Ertle, 2006; Simola, 2005; Romainville, 2002). הטענה היא שמנגנוני הערכה אחידים יכינו את התלמידים טוב יותר למערכת הכלכלה הגלובלית (Henry, Lingard, Rizvi and Talor, 2001). המשמעות המעשית של המבחנים האלה היא התעלמות ממורכבות תרבותית-חברתית וויתור על הכרה בשונות, ועל בחינת המורכבות התרבותית, החברתית והיסטורית, לטובת ההיגיון של הכלכלה הגלובלית (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). חוקרים אחרים מתארים את הריבוד הבין מדינתי והשלכותיו השליליות (Ertle, 2006; Simola, 2005; Romainville, 2002).

תחום אחר שבו המדינה מעוניינת בשליטה הוא תחום תוכניות הלימודים (Shohami, 2001; Smith, O'Day & Fuhrman, 1994). מדינת ישראל ידעה ויכוחים רבים בנושא, החל מביטול הזרמים בחינוך ועד לתוכנית הליבה. תוכניות לימודים גם הן נשלטות היטב באמצעות הערכה.

מדיניות כלפי הכשרת עובדי הוראה להערכה

כל חשיבה על מדיניות חינוכית נוגעת גם בעובדי ההוראה ומכאן – בהכשרתם. חשיבות איכותו של כוח ההוראה להשגת יעדים חברתיים ותרבותיים שונים, אינה מוטלת בספק, ומחשבה רבה מושקעת בארץ ובעולם, בכל מה שאמור בהכשרתם (Darling, 2003; Chochran-Smith, 2003; Hamond, 2000; Elliott, 1999; Freiman-Nemser & Remillard, 1996; Fullan, 1993; Fullan, 2001; Goodlad, 1995; Richardson, 1997).

כאשר מדובר במדיניות שעוסקת בהכשרת עובדי הוראה בישראל, ישנם מספר גופים שמשחקים תפקיד: 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 (נידרלנד, הופמן ודרור, בתהליך; בריטברד וטביביאן-מזרחי, 2003).

משרד החינוך, גורם משמעותי בקביעת מדיניות, מתייחס לסוגיה של הערכה בהקשר של הכשרת עובדי הוראה, פעם אחת בלבד, וזאת במה שאמור בהסמכתם. "קיום הליך תקין של הערכה הוא תנאי לקבלת רישוי לעיסוק בהוראה" (חוזר מנכ"ל, 2000). הערכה אינה מוזכרת בכלל במה שאמור בידע המקצועי של מורים, אלא כאשר דנים ברכז הערכה בית ספרי, שאז מדובר בתפקיד מובחן, ולא בידע מקצועי שוטף (חוזר מנכ"ל, 2008).

מספר ועדות חשובות, הן של משרד החינוך והן של המועצה להשכלה גבוהה, עסקו בסוגיה של עובדי הוראה והכשרתם. בין המוקדמות ניתן להזכיר את "הוועדה הציבורית לבעיות הכשרת המורים" (ועדת דושקין, 1962), שבראשה עמד עמנואל יפה, ושהחליטה על הרפורמה במבנה המוסדות להכשרת מורים והעמדת משך הלימודים על שלוש שנות לימוד, במקום שנתיים, ואת ועדת יפה (1971) שטבעה את רעיון האקדמיזציה: על המורה להיות בעל תואר שווה ערך לתואר אוניברסיטאי, ואילו מורי המורים חייבים להיות בעלי תואר שני לפחות. בעקבות ועדה זו נטבעה על ידי עמנואל יפה הדיכוטומיה "אקדמיזציה מול הומניזציה" (אורמיאן, 1979) ששרויה עד היום, באופן גלוי או סמוי, בשיח על הכשרת מורים.

ועדה בולטת אחרת שעסקה בנושא היא הוועדה בראשות השופט עציוני, שהוקמה ב-1979 על ידי שר החינוך דאז זבולון המר. ועדה זו המשיכה לקדם את רעיון האקדמיזציה, החל מקריטריונים למיון תלמידים (תעודת בגרות מלאה), וכלה בסטנדרטים אקדמאיים בלתי מתפשרים של סגל ההוראה במכללות.

המועצה להשכלה גבוהה מיהרה לאשרר את רעיון האקדמיזציה ב"דגם המנחה" שחובר בשנת 1981. דגם זה מחייב את המכללות ל"עמידה בסטנדרטים המצדיקים הענקת תואר אקדמאי" (ואן-גלדר, 2003; ועדת הקבע למסלולים אקדמאיים, 1981).

ועדת בן פרץ שהוקמה על ידי יוסי שריד, כבר קיבלה את האקדמיזציה כמובנת מאליה (נידרלנד, הופמן ודרור, 2009) ועסקה בעיקר בתשתיות, אבל גם ברמת המורים ומעמדם, ויצאה בדרישה להעלות את סף הדרישות לקבלה למסגרות הכשרת מורים, ולקיים בחינת רישוי ממלכתית. הוועדה גם המליצה לסגור חלק מהמכללות להכשרת עובדי הוראה (ועדת החינוך של הכנסת, 2002).

בכל הוועדות האלה שהתייחסו באופן נרחב לתהליכי הכשרת המורים, שצינו בכאב את ההישגים הנמוכים יחסית לציפיות של הסטודנטים להוראה, ושחשבו וניתחו לעומק מה הם

צריכים לדעת, לא מוזכרת ההערכה מכל סוג שהוא. לא הערכת התלמידים על ידי המורים, ולא הערכת המורים את עבודת עצמם. היכולת להעריך תוכניות לימודים או תוכניות של התערבות חינוכית – כל אלו היו לגמרי מחוץ לשדה הראייה, כך עד להקמת "כוח המשימה הלאומי לקידום החינוך" שברשות איש העסקים שלמה דוברת ("ועדת דוברת"), שאחזר אליה מאוחר יותר. מסקנות ועדת אריאב של המועצה להשכלה גבוהה, "מתווה אריאב", הן יוצאות דופן מכמה בחינות מהועדות שקדמו לה, הן ברוחב היריעה שהיא מטפלת בו, והן במתווה המדויק והמחייב שאיתו יצאה, מתווה שמחליף למעשה את "דגם המנחה" בן למעלה מעשרים שנה. ועדת אריאב משנה מעט את השיח שעוסק במתח שבין 'אקדמיזציה' ו'הומניזציה' ומדברת על מורה פרופסיונאלי (ועדת אריאב, 2006; נידרלנד, הופמן ודרור, בתהליך). ולמרות שהועדה 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 מופיעים כחלק מהמתווה. עם זאת, כולל המתווה כחלק ממתווה הבסיס (שם, לוח 1, עמ' 6) אוריינות מחקר בחינוך והוראה (לפחות 2 ש"ש) כאשר הכוונה ב"אוריינות מחקר" מפורטת מעט אחר כך "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 עמ' 7). עוד מוזכר במתווה, בסעיף קודם שעוסק בהוראה דיסציפלינארית, שבנוסף לרשימה ארוכה של נושאים שהמורה צריך להכיר בתחום הדיסציפלינה הוא צריך להכיר גם "גישות ושיטות הוראה **והערכה** המקובלות בדיסציפלינה" (שם, עמ' 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 והערכה בחינוך בפרט היא תחום מורכב ומבלבל הייתי רוצה לנסות ולעשות קצת סדר בנושא.

מהות ההערכה בחינוך

עולמם של המורים רווי בהערכה. מאז ראשית האנושות הערכה היא חלק מתהליך ההוראה (בירנבוים, 1997). מלמדים בכלל, ומורים בפרט, תמיד העריכו באופנים שונים את טיבה של עבודת תלמידיהם. אבל להערכה בחינוך גם פן אחר, הפן של ההערכה המערכתית ושל המדידה, ואלו הולכים ומתמקצעים מאז סוף המאה ה-19 ואילך. עבודות כמו של בינה, סטנפורד, גלטון, ספירמן ופירסון מחזירים למערכת מדידות רחבות היקף (Guba & Lincoln, 1989) שדורשות ידע סטטיסטי נרחב שלא היה ועדיין איננו מצוי בידי המורה מן השורה. ככל שהלכו תהליכי המדידה החינוכית והתפתחו, כך התרחקו מתהליך ההוראה של המורה בכתה, ומהישג ידו המקצועית. המדידה החינוכית נדדה מן החינוך אל הפסיכולוגיה וגם קיבלה שם חדש – פסיכומטריקה (לוי-רוזליס ולפידות, בתהליך).

מאז שטבע רלף טיילור את המושג 'הערכה' ('evaluation'), לא השתנתה הרבה הגדרתה הבסיסית - "קביעת ערכו או שוויו של דבר" (Tyler, 1942). מה שהשתנה, והרבה הוא היקף היריעה וכמות השאלות שההגדרה הפשוטה הזו מעוררת. לא נכון להתייחס להערכה בחינוך בלי להכיר לפחות חלק מהשאלות והשפעתן. התפישה הרווחת שהערכה זה טוב, היא אולי נכונה

בבסיסה. אבל עלינו להמשיך ולשאול את עצמנו שאלות כמו: טוב למי? באילו תנאים? מה המשמעות? מי מרוויח ומי מפסיד, ומה?

השאלות הגדולות הן כמובן מי מעריך? את מה הוא מעריך? איך הוא מעריך? לשם מה הוא מעריך? מה ייעשה עם התוצאות ואת מי הן תשמשנה? כל שאלה כזו מעוררת סוגיות נוספות.

מי מעריך? מעריך חיצוני או מעריך פנימי? (Nevo, 2001; Love, 1995, 2005) והמושג 'פנימי-חיצוני' הוא לעיתים עניין של נקודת ראות, והוא נוגע ברמות שונות, החל מרמת הכתה והערכת הישגי תלמידים, וכלה ברמת מוסד או אף מערכת (החל בתלמיד שמעריך את עצמו וכלה במבחנים בינלאומיים, חיצוניים למדינה). מומחה מתחום תוכן או מומחה בהערכה? (Levin-Rozalis, 2003) וזה מוביל לשאלות כמו מה ההכשרה שצריך מעריך? מה המיצוב ומה התפקיד של המעריך ושל ההערכה בתוך בית-ספר?

את מה מעריכים? תוכניות התערבות? מוסד חינוכי? מורה? מנהל? תלמיד? תוכנית לימודים? מדיניות חינוכית? חזון חינוכי? הישגי תלמידים?

איך מעריכים? האם מעריכים תהליך או מעריכים תוצר (Stake, 1967)? האם ההערכה היא מעצבת לשם שיפור, או מסכמת לשם 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 האלה נובעות מתפקידיה השונים ואף הסותרים של ההערכה בחינוך, שנובעים משתי תפישות מרכזיות, תפישה ששמה את המבנה במרכז מול תפישה ששמה במרכז את הפועל, או הסוכן האנושי (Berger & Luckmann, 1966). על פי תפישת המבנה, המערכת החברתית והמבנה שלה הם הגורם המשפיע על המציאות של היחידים, ורק אם נבין את המבנה נוכל להבין תופעות חברתיות שונות. גישת הסוכן האנושי או הפועל, תופשת את היחידים כמשפיעים על עולמם, ורק אם נבין את התהליכים של היחיד נוכל להשפיע. ניתן לארגן את תפקידי ההערכה בחינוך בתוך הדיכוטומיה הזו באופן שמוצג בלוח הבא (Levin-Rozalis, Rosenstein & Cousins, 2009 תרגום שלי מ.ל.ר.).

לוח מס' 1: תפקידי הערכה סותרים כנובעים מהבדלי תפישה

ממד	גישת ה"מבנה"	גישת "הסוכן האנושי"
פונקציה	פיקוח, בקרה ושליטה	למידה
מטרה	סטנדרטיזציה	הכרה בשונות, במגוון ובהבדלים ובחינתם
מסגרת	המבנה. מבט מערכתי	דיאגנוסטית (הבחנה בין תלמידים, מורים, או מוסדות)
מוקד	תוצרים	תהליכים (למשל: הערכה חלופית)
התוצאה	מיון	חיזוק, פיתוח וקידום
תוצר חינוכי	ידע	מיומנויות

מתודולוגיה	ניסויית	מגיבה
גישת חקר	אנליטית	הוליסטית
מיקוד	חיצוני	פנימי

מעניין לראות את ההקבלה שקיימת (אם כי לא המוחלטת) בין התפישות הפסיכומטריות, לבין הטור הימני (המבנה), ואת זו הקיימת בין הטור השמאלי (הסוכן האנושי) לעבודתו של המורה בכתה.

הדיון איננו דיון אקדמי של הגדרה כזו לעומת אחרת. יש לו השלכות מעשיות: לדרך שבה נעשית ההערכה, לשאלותיה, לכלים שלה, לקריטריונים ולסטנדרטים שלה, ליחסי הכוח שבין המערך למוערך ושבין יוזם ההערכה למוערך, וליחסי האמון ביניהם - יש השפעות מרחיקות לכת על הדרך שבה פועל ויפעל המוערך (מושא ההערכה) בין אם זה ארגון, מורה, תלמיד או מערכת שלמה. ולכן לכל בחירה שנעשית לגבי דפוס ההערכה במערכת החינוך ואופייה יש השפעות (שקשה להגזים בהן) על אופי מערכת החינוך, ובעיקר על דפוס הוראה ולמידה (בירנבוים, 2006; לוי, ללא תאריך; לוי-רוזליס ולפידות, בתהליך; Alkin, 1970; Bloom, 1971; Levin-Rozalis & Lapidot, in process).

להשלכות ולהשפעות הישירות או העקיפות של עצם השימוש בהערכה ומדידה; להשלכות של השימוש באופן מסוים של הערכה, לעומת אופן אחר; או בהשלכות חברתיות, ארגוניות או אישיות של עצם השימוש בהערכה, מדידה או בחינה קוראים 'תוקף השלכת' והוא נדון רבות בספרות החינוכית ובספרות הערכה (ראו למשל: Cuban, 2007; Greenwald et al, 2006; Doran, 2003; Jennings & Rentner, 2006; Jones and Olkin, 2004; Howe, 2005; Koretz, 2004; Messick et al, 1983; Messick, 1989; Patel & Russon, 1998; Sarason, 1998). תוקף השלכתי מתייחס למספר היבטים אישיים וחברתיים. ואין הכוונה לשימוש בממצאים, למידה מהם ושיפור, או על השפעה לטוב ולרע על מוערכים, אלא על השפעות "מערכתיות" גדולות, ובעיקר: (א) הבחינות עשויות להשפיע על תפישות של קובעי מדיניות, הציבור וגם של מתנכים לגבי מה הן סוגיות חינוכיות ראויות (ראו מיצ"ב ומבחנים בינ"ל), ובכך עשויים המבחנים לקבוע את סדר היום החינוכי או אף את תוכנית הלימודים. (ב) הקבוצה שתחבר את הבחינות תקבע במובנים רבים את הסטנדרטים החינוכיים. (ג) למערכת תהיה נטייה "ללמד לבחינה". זאת בלי להזכיר את ההונאות הקטנות והגדולות שצומחות סביב מבחנים ובעיקר קובעים (High stake).

הערכה בישראל, בתחום החינוך ומחוצה לו צמחה בעצלתיים בהשוואה לעולם המערבי. היו מעט תוכניות הכשרה, ופעולות ההערכה שנעשו, גם רחבות ההיקף שבהן, היו בעיקרן יוזמות של יחידים או עקב דרישות של קרנות חיצוניות (הלוי, 2009; Levin-Rozalis & Shochot-Reich, 2009; Schwartz, 1998).

במשרד החינוך נעשו עבודות הערכה באופן ספוראדי, אבל רק פעולה משמעותית אחת במה שאמור בהכשרת סגלי הוראה. ב-1993 חברו יחד משרד המדען הראשי, האגף לחינוך יסודי והאגף

לתכניות לימודים, לצורך פיתוח 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, עבור שתי דרגות כיתה עיקריות: ד' ו', אך שימשו בפועל ארבע דרגות כיתה ג'-ו'. במשרד הוקמה יחידת "המשוב הבית ספרי", שיועדה להטמיע את מאגרי המשימות ולעודד הערכה פנימית בבתי הספר. עם כניסת המיצ"ב דעך השימוש במשימות, והמאגר נמצא על אתר הראמ"ה (הרשות הארצית למדידה והערכה) לשימוש פתוח של מורים (מירו, 2009).

עם פרסום ההמלצות של כוח המשימה הלאומי לשיפור החינוך (התוכנית הלאומית לחינוך, 2005), השתנתה באחת ההתייחסות להערכה. כוח המשימה הלאומי יצא עם המלצה לשלב איש הערכה בכל אחד מבתי הספר בישראל. המלצה נוספת של כוח המשימה הלאומי היתה להקים גוף הערכה בלתי תלוי – רשות ארצית למדידה והערכה (ראמ"ה). לשתי ההחלטות יש השלכות מרחיקות לכת על תפקיד המורה ועל מערכת החינוך. ההחלטות האלה שמו את נושא הערכה בכלל, ובעיקר את תחום הערכה החינוכית, על סדר היום הציבורי.

הדיון הציבורי בהערכה בחינוך הוא היום רחב מאד ושותפים לו באמירה או במעשה גופים שונים כמו האקדמיה הלאומית הישראלית למדעים (באמצעות הועדה למחקר יישומי בחינוך ובשיתוף עם יד הנדיב), המועצה להשכלה גבוהה, משרד החינוך, מכשירי פרחי הוראה (מכללות ואוניברסיטאות), מכשירי מורים מכהנים (מכללות, מרכזי פסג"ה¹, אוניברסיטאות, גופים פרטיים ואחרים), איל"ת (האגודה הישראלית להערכת תוכניות), אפ"י (אגודה פסיכומטרית ישראלית). לכל אחד מהגופים תפישת הערכה משלו, סדר יום משלו, וכל אחד מהם מושך לכיוונו. מבחינת המדיניות להכשרת עובדי הוראה להערכה, זו לא השתנתה. השיח הער והמשמעותי כמו פסח עליה, למרות שכפי שאראה מייד, היו גופים בעלי השפעה שהכניסו את המורים לשיח שהחל מתפתח בנושא.

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

תפישות מרכזיות בשיח על הערכה חינוכית בישראל

כוח המשימה הלאומי לשיפור החינוך (ועדת דוברת):

התפישה שבאה לביטוי בהמלצות כוח המשימה הלאומי (התוכנית הלאומית לחינוך, 2005), היא בעיקרה של ניהול מבוסס תוצאות (evidence based management), שמשמעותו, כפי שבאה לידי ביטוי בהמלצות הועדה, אחריותיות (accountability) ובקרה. כך בהגדרת הבעיה על ידי הועדה: *"... תרבות קבלת ההחלטות אינה נשענת באופן הנדרש על קבלת נתונים וניתוחם המקצועי ... תוכניות לימודים רבות אומצו בעבר על סמך עמדות (ואופנות), ולא התבצעו תהליכים מסודרים של מדידה והערכה לבדיקת יעילותם..."* (התוכנית הלאומית לחינוך, 2005, עמ' 29).

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*

¹ פסג"ה – פיתוח סגלי הוראה, גף במשרד החינוך שאחראי לנושא של השתלמויות מורים מכהנים. הפסג"ה מפוזרות בכל חלקי הארץ.

תנאים הכרחיים למימוש עיקרון האחידות בכל דרגיה של מערכת החינוך... (שם, עמ' 37). בהמשך מסבירה הועדה את המשמעות של אחריותיות: "...על-פי עקרון האחידותיות (Accountability), כל דרג חופשי מהתערבות מיותרת בפעולתו מצד הדרג הממונה עליו, אולם חייב במדידה ובהערכה של פעולותיו והישגיו תוך שקיפות מלאה. ... משנקבעו היעדים, לבקר ולוודא כי הפעילות נעשית בהתאם ליעדים ולכללים שנקבעו..." (שם, עמ' 42).

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

על מנת לממש את התפישה הזו הוחלט כאמור על הקמת הראמ"ה (שם, עמ' 74), ועל שילוב איש הערכה בכל בית ספר (שם, עמ' 54). ואכן, אחת התקנות שתיקן משרד החינוך בעקבות הפעילות בשנת 2004 של כוח המשימה הלאומי (ועדת דוברת), הפכה את גף הערכה לרשות ארצית למדידה והערכה (ראמ"ה), גוף סמך אוטונומי שאחראי הן להערכה במערכת החינוך והן להכשרת המעריכים.

ההחלטה על שילוב איש הערכה בכל בית ספר וכתוצאה מכך הצורך להכשיר אלפי מעריכים עבור מערכת החינוך יצר בישראל מערבולת של הכשרות וניירות עמדה, חלקם ראויים ורובם לא (לוי-רוזליס, לפידות ודובר, 2006; לוי-רוזליס ולפידות, 2008).

משרד החינוך, והועדה למחקר יישומי בחינוך שליד המועצה הלאומית הישראלית למדעים, הזדרזו להגיב להחלטות של כוח המשימה הלאומי בהקשר של הערכה. כל אחד מהם יצא עם מסמך משלו, שמבטא השקפת עולם ומתווה דרכי פעולה. משהוקמה הראמ"ה יצאה גם היא עם מסמך משלה. ברצוני להציג את שלושת המסמכים האלה שהם במרכז השיח סביב 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. הועדה פרסמה את מסקנותיה במסמך שהוגש למנכ"לית המשרד (ועדת היישום להערכה פנימית, 2004).

הועדה ניסתה לעמוד בפרץ של ההחלטה הגורפת לשלב איש הערכה בכל אחד מבתי הספר בישראל, שלא התייחסה כלל לסוגיית הכשרתם המקצועית של אלפי המעריכים שידרשו למערכת. הוועדה הדגישה את החשיבות העצומה של הכשרה מקצועית ראויה של איש הערכה הפנימי. המלצה שאומצה מאוחר יותר, עם ההחלטה שרכז הערכה בית ספרי צריך להיות בעל תואר שני ייעודי בהערכה (חוזר מנכ"ל, 2008).

נקודת המוצא של ועדת היישום מקבלת את הנחות כוח המשימה הלאומי אבל נותנת להם דגשים משלה (ההדגשות בטקסט שלי. מ.ל.ר.):

“ קבלת החלטות חייבת להתבסס על נתונים אמין 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
- נעשתה 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, עמ’ 1).

הסוגיה של מקצועיות מודגשת מאד במסמך ולכל אורכו. במסמך בן עשרה עמודים מופיעה המילה **מקצועי** על נגזרותיה, בהקשר של עבודת רכז ההערכה, חמישים פעמים. הוועדה גם מגדירה מה היא המקצועיות הנדרשת, הן מבחינת תכניה, והן מהבחינה הפורמאלית. מבחינת התכנים מגדירה הוועדה את הנושאים הבאים: **“ כדי שלהערכה תהיה תרומה משמעותית חייב המעריך להיות איש מקצוע, השולט בידע ובשיטות הקיימים במקצוע לדבות המיומנויות הבאות (ראה פירוט נוסף בנספח ב’):**

- יכולת לשותף 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, עמ’ 1).

מבחינה פורמאלית מגדירה הוועדה את רף ההכשרה המקצועית הנדרשת: **“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. כמו-כך יועצים משמשים רק אנשים שהוסמכו לכך על-ידי מוסד להשכלה גבוהה והמנחה המקצועי שלהם הוא שפ”י” (שם, עמ’ 3). ובהמשך: “הוועדה ממליצה על הקמת מסלול אקדמי להכשרת מעריכים שיהיו בעלי תואר שני עם התמחות בהערכה” (שם, עמ’ 6). המלצה שכאמור התקבלה.**

הוועדה מודעת לקושי להכשיר אלפי רכזי הערכה בעלי תואר שני ומציעה שלב ביניים של הכשרה בת 112 שעות, למה שמכונה **“רכז מידע ומשוב”**. המיומנויות הנדרשות מרכז המידע הן רבות ומורכבות גם כן, שעיקרן יכולת גבוהה מאד להשתמש במידע ממנב”ס וממאגרים ארציים, הפעלה עקבית ושיטתית של כלי הערכה פנימיים, מקצועיים ובדוקים, הקיימים כבר במערכת החינוך,

ועיבוד המידע ממקורות פנימיים וחיצוניים כך שיוכל לקדם את בית-הספר בצעד גדול בפיתוח תרבות הערכה, ותהליכי קבלת החלטות מבוססות נתונים (שם, עמ' 4).

ועדת היישום עושה הבחנה חשובה בין הערכה פנימית וחיצונית (עמ' 2-3), מדגישה את חשיבות המערך המקצועי התומך לאנשי ההערכה הבית ספריים (עמ' 7), ומציינת את כלי ההערכה המקצועיים שעומדים היום במערכת לרשות המעריכים.

הוועדה מדגישה מאד את רמתם המקצועית הגבוהה והכשרתם המקצועית הייעודית של אנשי ההערכה הבית ספריים, אבל מחברת את עבודתם לבית הספר, ולצרכיו, מדגישה את חשיבותו של ידע פדגוגי ושל בית הספר כצרכן המרכזי של רכז ההערכה. בהתייחס להחלטות כוח המשימה הלאומי, מאמצת וועדת היישום את ההערכה ככלי ניהולי, אבל כלי ניהולי פנימי של בית הספר לשם קידומו, ולא כלי פיקוח מערכתי כפי שדרשו החלטות כוח המשימה הלאומי.

האקדמיה הישראלית הלאומית למדעים

היוזמה למחקר יישומי בחינוך, שליד האקדמיה הישראלית הלאומית למדעים ובשיתוף יד הנדיב, נכנסה גם היא לשיח סביב ההערכה, במספר פעולות משמעותיות. בשלב הראשון נערכו שני ימי עיון רחבי היקף (בדצמבר 2004 ובינואר 2005) שהביאו לארץ אנשי מדידה חינוכית בעלי שם בינלאומי (האקדמיה הלאומית הישראלית למדעים, 2005). בעקבות ימי העיון מינתה ועדת מומחים למדידה ולהערכה בחינוך, שדנה במה צריכים לדעת העוסקים במדידה ובהערכה בחינוך. הוועדה הוציאה מסמך "הצעה למסגרת לתכניות לימודים ולפיתוח מקצועי" מסגרת לתוכנית לימודים, שהוגש למועצה להשכלה גבוהה (מ"ג), לוועדה לתקצוב ותכנון (ות"ת), למשרד החינוך, ללשכת המדען הראשי, והרשות הארצית למדידה והערכה (הראמ"ה), ולמוסדות אקדמאיים (הוועדה למדידה ולהערכה בחינוך, 2005). בקיץ 2006 קיימה היוזמה קורס קיץ מרוכז בנושא: "Evaluating Effectiveness in Education" בראשות פרופסור ברוש מאוניברסיטת פנסילבניה. הקורס נועד למי שמלמדים את הנושא במוסדות אקדמיים, לאנשי הראמ"ה ומשרד החינוך ולמעריכים ממוסדות שונים.

מסמך "הצעה למסגרת לתכניות לימודים ולפיתוח מקצועי" של הוועדה למדידה ולהערכה בחינוך הוא מסמך מקיף ומפורט, שעוסק באוכלוסיות השונות במערכת החינוך, ובסוגי הידע ההצהרתי 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, עמ' 5). הבקרה והאחריות לא נמצאים במסמך, ואת הנחת היסוד השלישית שלו ניתן 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, עמ' 5).

חברי הוועדה גם מזהירים מפני התוקף ההשלכתי, ומשימוש לא מבוקר בהערכה: "לא רק שמוכן שעל משרד החינוך להתוות את המטרות והיעדים למערכת החינוך, אלא שיש להיזהר ממצב בו בהעדר מטרות שהוגדרו על ידי האחראים, הופכים המודדים והמעריכים לקובעי מטרות בפועל. בהקשר זה חשוב להזכיר, שידוע שהיחס שבין מדידה והערכה מחד, לבין 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 אודות הקשרים בין מטרות חינוכיות שונות לבין שיטות שונות של הערכה" (שם, עמ' 5).

ה"מסגרת לתוכנית לימודים" עושה מספר הבחנות חשובות בין הערכה חינוכית להערכת תוכניות ופרויקטים, ובתוך הערכה חינוכית בין הערכה לצורך למידה (ה"ל") והערכה של הלמידה (הש"ל) על מטרותיהן השונות. המסמך הזה הוא הראשון שמתייחס להבדלים בין הערכה חינוכית כחלק מעבודת ההוראה, לבין מדידה חינוכית ופסיכומטריקה. כנראה לכן הוא גם הראשון שמתייחס להכשרתם בתחום ההערכה של המורים בשדה: "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 לעיתים לבודד את תפקיד המורה כמעריך משאר תפקידיו" (שם, עמ' 27). המסמך מציע שלושה קורסי בסיס: קורס מבוא (שמקנה מושגי יסוד וידע בגישות להערכה, תפקידיה, שיקולים שונים בביצוע ההערכה, והקשר בינה לבין משתנים חינוכיים שונים); קורס בסיסי בסטטיסטיקה ומדידה; וקורס שעוסק בהערכה בכיתה (classroom assessment) (שם, עמ' 17 ו-27).

הרשות הארצית למדידה והערכה

כאמור, ההחלטה על הקמת הראמ"ה יצאה לפועל גם היא. הראמ"ה אכן הוקמה, ופעילותה מורגשת. אחד הדברים המשמעותיים שעשתה הראמ"ה בהקשר הזה הוא קביעת סטנדרטים לידע מקצועי של מעריכים (הרטף, גנור, רום ושלטון, 2007). חוברת המפרטים מציגת ש: "תחום ההערכה הפך להיות תחום מרכזי בתהליך החינוכי. תפקיד ההערכה במונח זה נתפס בזיקה גבוהה לתהליך ההוראה-למידה וכחלק בלתי נפרד ממנו ומחייב ידע ומיומנויות מקצועיות מתקדמות" (הרטף, גנור, רום ושלטון, 2007, עמ' 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, עמ' 1). החוברת שמתבססת על שני המסמכים הקודמים (ועדת היישום, והוועדה למדידה והערכה בחינוך) מפורטת מאד. היא מבחינה בין אוכלוסיות יעד שונות וצורכיהן, בין סוגים של גישות שונות להערכה, מדרגת תחומי ליבה ותחומים בדרגת חשיבות שנייה.

בשונה משני המסמכים הקודמים, המסמך הזה איננו רק הצעה, ובתחומים מסוימים יש לו שינויים. כך למשל, במה שאמור במרכזי הפסג"ה. אם בשנת 2004 התקיימו 77 השתלמויות בהערכה במרכזי הפסג"ה, ובשנת 2006 שמונים ושתיים השתלמויות כאלה, הרי בשנה שלאחר יציאת המפרטים, והפיקוח על תכני ובעיקר רמת ההשתלמויות מצאנו רק ארבע השתלמויות (לוי-רוזליס ולפידות, 2007). השתלמויות שלא עמדו בקריטריונים לא אושרו.

לסיכום

בין ארבעת המסמכים: התוכנית הלאומית לחינוך (כוח המשימה הלאומי), המלצות ועדת היישום להערכה פנימית (משרד החינוך, במסגרת כוח המשימה הלאומי), הצעה למסגרת לתוכניות לימודים ולפיתוח מקצועי (היוזמה למחקר יישומי בחינוך), וחברת המפרטים (הראמ"ה) ניתן לראות התפתחות בתפישת מורכבות התחום. אם החלטות כוח המשימה הלאומי היו גורפות ולא נכנסו לפרטים (הקמת הראמ"ה ושילוב מעריך פנימי), הרי שוועדת היישום חשבה על ה"איך": מהי הדרך הראויה לבצע את ההחלטה. שני המסמכים הבאים, מסמך הוועדה למדידה והערכה, וחברת המפרטים של הראמ"ה חשבו גם על ה"מי", ה"מה" וה"כמה": מי הם קהלי היעד, מה הם צריכים לדעת וכמה. כאשר רמת הפירוט של חברת המפרטים רבה יותר משל מסמך הוועדה למדידה והערכה.

יש גם שינוי בהשקפת העולם, החל מתפישה נוקשה של אחריותיות (accountability) ב'תוכנית הלאומית', לתפישה של למידה מהערכה של כל מרכיבי המערכת, כולל למידה עצמית ב'חברת המפרטים'.

במה שאמור בהכשרה להערכה של סגלי ההוראה – שני המסמכים האחרונים רואים חשיבות רבה בהכשרת המורים ומגדירים בפירוט תחומי דעת שאלו צריכים לדעת, הן כחלק אינטגרלי מתפקידם כמורים ועוד יותר בעידן שבו תהליכי הערכה במערכת החינוך הולכים ומתפתחים. חשוב עוד לציין שהשיח סביב ההערכה לא מסתיים במסמכים הכתובים. היוזמה למחקר יישומי בחינוך עשתה עוד מספר פעולות שקשורות לנושא, שהבולטת שבהן היא הנחיה להערכה ומדידה בגיל הרך (ברזניץ וימין, 2008). הראמ"ה, במקביל לפעולותיה הרבות שבתוקף תפקידה כרשות ארצית למדידה והערכה, פועלת רבות להעלאת אוריינות ההערכה במערכת החינוך. הרשות מקיימת ימי עיון רבים לקבוצות אוכלוסייה שונות, השתלמויות שונות, מביאה אנשי שם מחו"ל, נציגיה משתתפים בכנסים וימי עיון שנעשים על ידי גופים אחרים, והאתר שבבעלותה מציג חומרים רבים ומגוונים לרשות הציבור.

שתי מערכות בעיקר אמונות על הכשרת סגלי ההוראה. המכללות להכשרת עובדי הוראה ומרכזי הפסג"ה. במקביל מקנות האוניברסיטאות תעודת הוראה. ישנם גופים שונים נוספים שיוזמים השתלמויות למורים. על פי נתוני אגף ההשתלמויות במינהל להכשרה, להשתלמויות ולהדרכה לעובדי הוראה במשרד החינוך, התקיימו בשנת 2005 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.

אציג כעת ממצאי מחקר שנערך בשנים 2006-2008 ובדק את הפעילויות הראשונות שנעשו בהקשר של הכשרה להערכה של סגלי הוראה.

מרכזי הפסג"ה

מרכזי הפסג"ה הגיבו מהר בריבוי השתלמויות למורים מכהנים. מממצאי המחקר שבדק הכשרה להערכה במרכזי הפסגה (לויין-רוזליס ולפידות, 2007) מראים שבשנת 2007 הפעילו מרכזי הפסג"ה 84 השתלמויות שונות בהערכה בכל רחבי הארץ. מנהלי מרכזי פסג"ה תופשים את השתלמויות ההערכה כחשובות ונחוצות ומגלים מעורבות רבה בכל הקורה בהשתלמות, יחס רציני וציפיות גבוהות מן ההשתלמות.

תכניות הלימודים מגוונות, מאוד. כוללות תערובת של יחידות מידע. כל השתלמות היא משהו אחר. מתוך 84 השתלמויות שונות הקשורות בהערכה שניתנו במרכזי הפסג"ה השונים ברחבי הארץ, זכו 8 השתלמויות לשם הכללי "הערכה" או "הערכה ומדידה" עשרים ותשע אחרות כרכו בתוכן למידה והערכה ("דרכי הוראה והערכה" הוא השם הנפוץ). ארבע השתלמויות יועדו לרכזי / נאמני הערכה. כל השאר (43 השתלמויות) עסקו בנושאים שונים כמו "הערכה בגן הילדים", "הערכה בלימוד מדעי כדור הארץ", "תרבות הערכה", "אקלים בית ספרי בדגש על הערכה", ועוד נושאים רבים ומגוונים.

המרצים המלמדים בהשתלמויות הם בעלי רקע דיסציפלינארי מגוון. רובם ככולם "אנשי הערכה" העוסקים בהערכה ומכירים בחשיבותה. הם מבקשים לראות במשתלמים סוכני שינוי, סוכני הטמעה של תרבות ההערכה הבית ספרית.

יוזמי ההשתלמויות מגוונים גם הם. היוזמה ל-39 השתלמויות מתוך 84 הייתה של מרכזי הפסג"ה עצמם. מפמ"רים, יחידות מטה שונות, וגופים אחרים בתוך משרד החינוך ומחוצה לו יזמו את השאר.

מעבירי ההשתלמויות בחלקם הם גופים כמו הטכניון, האוניברסיטה הפתוחה, מט"ח, אוניברסיטת בר אילן, אוניברסיטת תל אביב, מכון ויצמן ומכללת אחוה. חלק אחר הם הגופים והאגפים בתוך משרד החינוך שיזמו את ההשתלמויות.

מתוך הממצאים נראה שתרומת ההשתלמויות (באופן שונה כמובן מאחת לשנייה) היא חלקית, וזאת מכמה סיבות - חלקן טכניות וחלקן מהותיות.

הבעיות הטכניות מתמצות במספר השעות הקטן שמוקדש לנושא שהוא כל כך רחב ומורכב. מספר השעות המועט גורם לרדידות של ההכשרה בתחום - השתלמויות קצרות, שמכסות מעט, ונבנות בהתאם לתפישת המרצה מה חשוב ומה לא, וזאת עבור ציבור משתלמים שהוא כמעט לחלוטין חסר רקע בתחום.

במקביל, ומכיוון שההערכה מחוברת כמעט לכל תחום בעשייה החינוכית, ישנם הרבה מאד גופים במערכת שיוזמים השתלמויות למורים בתחום ספציפי ממוקד של הערכה. יש לכך כמובן יתרונות רבים, אבל במבט כללי נוצרת פרגמנטציה של התחום, ובלבול גדול, כאשר לא תמיד ברור על מה בדיוק מדברים, לאן הידע המסוים שייך, לאיזה תחום, לאיזו גישה, ומתוך איזו השקפת עולם.

מנהלי מרכזי הפסג"ה הגיבו לצורך המזוהה בהכשרה להערכה, וגם נתנו גג ובית לגופים מכשירים חיצוניים. אולם בהיעדר מדיניות ברורה לגבי מה לומדים ולצורך מה, נראה כל מערך ההכשרה די כאוטי (לויין-רוזליס ולפידות, 2007). את מקום המדיניות החסרה תפסה חוברת המפרטים של

הראמ"ה שמנסה לתת תשובות מסודרות לכל השאלות האלה, וגם יש לה את הכוח האדמיניסטרטיבי לקבוע נורמות וכללים להערכה ראוייה.

מדיניות ההכשרה להערכה במכללות להכשרת עובדי הוראה

הכשרה להערכה של פרחי הוראה במכללות להכשרת עובדי הוראה ובאוניברסיטאות, נבדקה ב-2006 ושוב ב-2008 (לוי-רוזליס ולפידות, 2006, 2008). הממצאים שעוסקים בהכשרה להערכה של פרחי הוראה רבים ומגוונים (פירוט בדו"חות המלאים), אבל מדיניות ברורה לא מצאנו.

ההכשרה להערכה לפרחי הוראה נעשית באופן מגוון ושונה מאוד בין מכללה למכללה. היא פרושה על פני סוגים שונים של תכניות לימוד ופונה לקהל מגוון. באופן כללי נחלקות התכניות לשתי קבוצות עיקריות: אחת, תכניות המיועדות לפרחי הוראה, קורסים ייעודיים להערכה, קורסים תלויי דיסציפלינה, במסגרת לימודי הפדגוגיה והחינוך, ובמסגרת סמינריונים דידקטיים ומחקריים. ושנייה, תכניות הנכללות בתוך מסגרת לימודי ההמשך: השתלמויות למורים ומנהלים מכהנים ובמסלולי התואר השני.

היקף הקורסים הנלמדים הוא מצומצם. רובם סימסטריליים ואינם מלווים בתרגול. מצד שני הם עשירים בתכנים כמו 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 וחיצונית, וכדומה. בחלק מהקורסים ניתן היה למצוא את כל הנושאים האלה, כאשר לכל אחד מהם מוקדש שיעור אחד.

המרצים שמלמדים הערכה מגיעים מדיסציפלינות שונות ולרובם אין הכשרה בהערכה או במחקר. כולם חושבים שהנושא חשוב, ולרובם ידע חלקי מאד ומבולבל באשר למהות התחום.

במכללות שבדקנו, לא מצאנו החלטה על מדיניות, חשיבה על גוף ידע ברור, על סדר עדיפויות לגבי מה חשוב ללמד ומה לא, מתי וכמה. במכון מופ"ת היתה קבוצת חשיבה וולונטרית, שמסקנותיה פורסמו בירחון מופ"ת, אבל עדיין לא ראינו השפעה שלה בשטח.

לתוכנית הלאומית לחינוך (דו"ח דוברת) היתה השפעה מסוימת על כמות הקורסים הניתנים לפרחי הוראה. בשנה הראשונה של המחקר (2006) מצאנו כ-20 קורסים ייעודיים וכ-30 קורסים להערכה תלויי דיסציפלינה, בחמש עשרה מכללות. בבדיקה חוזרת שערכנו בשנת המחקר השלישית (2008) מצאנו 28 קורסים ייעודיים להערכה ו-40 קורסים תלויי דיסציפלינה ב-18 מכללות. מספר הקורסים בתואר השני כמעט והכפיל את עצמו, וגדל מתשעה קורסים בשנת המחקר הראשונה ל-17 בשלישית. לא נמצאו הבדלים בתכנים.

מדיניות ההכשרה להערכה באוניברסיטאות

המצב באוניברסיטאות הוא דומה לזה שבמכללות למעט דבר אחד. המרצים שמלמדים הערכה באים מהתחום, והתפישה שלהם את נושא ההערכה ואת רציונל מבנה הקורס שהם מלמדים קוהרנטית ומדויקת יותר. אבל גם כאן ההחלטה על מבנה הקורס ותכניו היא בידי מלמד הקורס.

לסיכום

הערכה איננה "דיסציפלינה" ולכן אין לה גוף ידע קנוני ברור. אם נוסיף לכך את תפקידיה השונים של הערכה בכלל והערכה בחינוך בפרט, ואת הגישות השונות הקיימות להערכה, נבין את הבלבול הרב השורר בתחום ההכשרה להערכה של עובדי הוראה. היעדר מדיניות ברורה שמתווה דרך לגבי מקומה של ההערכה בתפקידם ולכן בהכשרתם של עובדי הוראה גוררת אחריה ארבע בעיות מרכזיות במה שאמור בהכשרתם להערכה של עובדי הוראה:

1. אין החלטה או חשיבה על גוף ידע קנוני, לפיכך גוף הידע לא ברור, ויש בלבול וחוסר הבחנה בין תפקידים שונים של הערכה
2. מחסור בכוח אדם מיומן
3. מיעוט זמן המוקדש להכשרה להערכה
4. העדר תכנית הכשרה מובנית

דיון - הסוגיות המרכזיות בנושא הערכה שקיימות היום בשיח סביב הערכה והשפעתן על עובדי ההוראה

דו"ח דוברת תרם להעלאת נושא ההערכה למרכז הבמה החינוכית והיפנה את הקשב והשיח הציבורי כלפי נחיצותה של ההערכה במערכת החינוך. חשוב לציין כי נושא ההערכה היה קיים במערכת קודם לכן, אך הדו"ח והמלצותיו היוו גורם מזמן לדיון ובירור נושא ההערכה בחינוך. מספר סוגיות ומתחים בלטו מלכתחילה במה שאמור בתפישת תפקידיה של ההערכה:

1. המתח שבין הערכה כתהליך בקרה ופיקוח לבין הערכה לצורך למידה
2. המתח שבין הערכה לפי מודלים מבניים קשיחים לבין הערכה מגיבה ותלויה שדה
3. הערכה של תוצרים לעומת הערכה של תהליכים ומורכבויות

המתח שבין הערכה כתהליך בקרה ופיקוח לבין הערכה לצורך למידה

הסוגיה הראשונה שעלתה על סדר היום, והדיה נשמעו בימי עיון ובכתבות עיתונאיות היתה סוגיית הפיקוח והאחריותיות. כפי שהראית, התפישה שממנה יצא כוח המשימה הלאומי (ועדת דוברת) היתה באופן גורף תפישה של אחריותיות (accountability) ופיקוח.

המקור של המושג ושל התפישה העומדת מאחוריו הוא בתיאוריות וגישות של ניהול הירארכי לפיהן הכפוף מוסר דין וחשבון לממונה עליו. זה גם השימוש היום כפי שעבר להערכה, בעיקר של הזרם שדורש שינוי במערכת החינוך מתוך המטפורה של "מפעל": תשומות ותפוקות מדידות, ברורות וחד משמעיות עם קשר חד-חד ערכי ביניהן וללא גורמים מתערבים, ללא השפעות ארוכות טווח, ללא תהליכי ביניים.

הרעיון של אחריותיות קיבל משנה תוקף ב-2001 עם החוק בארה"ב של No Child Left Behind (NCLB) – The reauthorization of the Elementary and Secondary Education Act, 2001. מאז תפס הנושא תאוצה, וכך גם הויכוח בעקבותיו (Hamilton, Stecher & Klein, 2002). אחת

הבעיות המרכזיות היא אי הבהירות של המושג, הבלבול בינו לבין 'אחריות', והשימוש השונה שעושים בו מקבלי החלטות שונים (Darling-Hammond & Snyder, 1992; Heim, 1995). פרידמן (2005) מונה שש דרגות של ביטויי אחריותיות וההשלכות שלהן שהולכות ומחמירות:

1. דיווח על ביצוע, באמצעות ממצאים מדידים. הבעיה: תהליכים, מורכבויות, תופעות לא מדידות יוצאות מהתמונה.
2. פיקוח ועמידה בסטנדרטים – בצעת או לא בצעת, זו חזות הכל. הבעיה: הרלוונטיות של הסטנדרטים, דחיפה של אנשים לא לקחת סיכונים בניסיונות שסיכויי ההצלחה שלהם נמוכים (עבודה עם תלמידים מתקשים, למשל).
3. תמריצים – תגמול על תוצאות (מחייבים את השלבים הקודמים: 1+2) – בשלב הזה מתחיל תגמול דיפרנציאלי על פי מידת העמידה או אי העמידה בסטנדרטים.
4. כוחות השוק – בשלב הזה מתחילים להתערב עיתונות, לחצים פוליטיים שונים, רלוונטיים ולא רלוונטיים שבאים לעיתים מתוך אינטרסים זרים, והרבה פעמים מתוך אי הבנה של העניין. סטנדרטים פופוליסטיים לעיתים מזיקים, נקבעים כתוצאה של הלחצים האלה.
5. בקרה חיצונית והקטנת אחריות עצמית – כתוצאה מכל מערכת הלחצים הזו נגמרת האחריות האישית לביצוע מקצועי, והמוערכים עושים הכל בשביל להתאים עצמם לסטנדרטים. אלו שלא מוכנים לכך נפלטים מהמערכת.
6. הגברת הלחץ של הפיקוח הפנימי – האווירה בתוך הארגונים לחוצה. יש לחץ של כולם על כולם לעמוד בסטנדרטים, מתחילה תחרות, ונעלם שיתוף פעולה שאינו מתגמל ואף נענשים עליו.

התפישה הקיצונית, כפי שבוטאה בתוכנית הלאומית לחינוך, 2005 אומנם לא נשמעה יותר. אבל סוגית האחריותיות עדיין מלווה את השיח סביב ההערכה. מן הצד השני ישנה התפישה שמדברת על הערכה כאמצעי ללמידה. הראשות הארצית למדידה והערכה (הראמ"ה) גם אמצה את הניב "מדידה לצורך למידה" כמוטו שלה.

למידה יכולה להיות של תלמיד או מורה, של מוסד וגם של המערכת כולה. ההבחנה שאנו עושים בין הערכה שמשמשת ללמידה לבין הערכה שמשמשת לפיקוח היא באמצעות "מבחן הנתונים". את מי הנתונים אמורים לשרת? האם את מושא הערכה (תלמיד, מורה, מוסד, תוכנית, מערכת) או את מי שהוא מחוץ למושא הערכה ויש לו השפעה על גורלו? (מורה מקבל את נתוני התלמידים, מנהל מקבל את נתוני המורים, וכן הלאה). כאשר מי שמזמין (ומקבל) את הנתונים איננו מושא הערכה אלא מי שמחוץ לה - הרי שמדובר למעשה בבקרה - גם אם בחלק מהמקרים מושא הערכה מקבל גם הוא את הנתונים.

למידה אמיתית תתרחש בעיקר כאשר שאלות הערכה מנוסחות יחד עם מושא הערכה ומותאמות לצרכיו. בדרך זו לומד מושא הערכה לשאול את עצמו שאלות לגבי עצמו, התפקוד שלו, הישגיו, ויכולתו להגדיר מטרות ולפעול ליישמן. כאשר שאלות אלה מוחלטות בשבילו, אם בכלל תהיה למידה - היא תהיה חלקית, וחלק מהלמידה יהיה כיצד להתחמק או כיצד "לצאת טוב" על פי הקריטריונים שקבע מישהו אחר, ולא כיצד לשפר את הטעון שיפור.

המתח שבין הערכה במודלים מבניים קשיחים לבין הערכה מגיבה ותלויה שדה

בהערכה במודלים מבניים קשיחים הכוונה למבחנים שמטרתם לצייר תמונה מערכתית של המוערך: כתיב, בי"ס או מבחנים רחבי היקף, שניתנים לכלל המערכת או באופן מדגמי. מדובר בבדיקה בעלת רמת אמינות גבוהה, וחשוב יותר, בעלת יכולת הכללה גבוהה. אחד המודלים המבניים הקיצוניים הוא המודל הניסויי (הקצאה אקראית לקבוצת ניסוי וקבוצת ביקורת – RCT: randomized controlled trials), שלמרות יתרונותיו שנוי מאד במחלוקת מכיוון שמשמעות הדבר היא שאופן הבדיקה מכתוב את אופן ההתערבות. תהליך הבדיקה חשוב יותר מאשר שיקולים מקצועיים (למי ואיך להקצות) או אתיים (קיפוח אלה שלא הוקצה להם). אבל גם ללא המודל הניסויי, למודלים המבניים הקשיחים ישנם חסרונות שעיקרם הוא חוסר רגישותם לתהליכים ולשונויות.

בקצה השני נמצאות המדידה וההערכה המגיבות, תלויות השדה. אלה שבדקות את המוערך באופן תהליכי, כחלק מההקשר שבו הוא מצוי ותוך שיתוף פעולה ודיאלוג בין המעריכים למוערכים. כלומר הבדיקה מתאימה את עצמה, את כלי הבדיקה, את שיטות הבדיקה, ואת שיטות ניתוח הנתונים לתנאי השדה, ההתערבות, האוכלוסייה וכדומה, ותופרת בדיקה "מעשה חייט" למידות המוערך.

למען הסר ספק ולמנוע טעות, בדיקה מגיבה איננה בהכרח איכותנית, היא יכולה להיעשות באופן כמותי, באמצעות שאלונים, מבחנים או כלי מדידה אחרים כאשר הממצאים מעובדים ומנותחים בשיטות סטטיסטיות מוקפדות. כמו כן הערכה במערך ניסויי איננה בהכרח כמותית: היא יכולה להשתמש בכלים פתוחים כמו ראיונות ותצפיות או ניתוח תוכן של מסמכים או מבחנים, למרות שהעיבוד הסופי ייעשה בסופו של התהליך במספרים ובכלים סטטיסטיים, שכן במערך זה אין דרך אחרת להכריע על טיב המוערך.

דפוס הבדיקה שישלוט במערכת החינוך יכתוב במידה רבה את אופי הפעולה החינוכית, ואת 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 האם גם תופעות ותהליכים שהם מקומיים, ייחודיים, הוליסטיים, נטורליסטיים ועמומים יותר יוכלו לשחק תפקיד בהבנת השדה החינוכי).

ישנו גורם נוסף שמבחין בין שני הקצוות שהוצגו כאן, וזו סוגיית המקצועיות של המערך. המודלים המבניים הקשיחים מצריכים רמת מקצועיות גבוהה בכל מה שאמור בבניית הכלים, תכנון הבדיקה, המדגמים, הניתוחים הסטטיסטיים והבנת משמעותם. את פעולות ההערכה האלה לא יכולים לבצע מורים. אבל מורים צריכים להיות מסוגלים לקרוא דו"חות הערכה כאלה ולהבין אותם, בעיקר אם התפישה המובילה היא של הערכה ומדידה לצורך למידה.

גם לא כל מה שאמור בהערכה מגיבה יכולים המורים לבצע. אבל הם יכולים להיות שותפים לדיאלוג לגבי מה מעריכים ולמה. מה הן שאלות ההערכה או המדידה. והם יכולים כמובן לתת למעריכים הרבה מידע הקשור במוערך (לויין-רוזליס, 2005). מורים כן יכולים לבצע, עם הנחיה ובלעדיה, הערכת תלמידים בגישות חלופיות, הערכה עצמית והערכת עמיתים. בתנאי שנושאים אלה יהיו חלק מהכשרתם (הוועדה למדידה ולהערכה בחינוך, 2005; הרטף וחוב' 2007).

המתח בין הערכה עם דגש על תוצרים בלבד, לבין מדידה והערכה שמתייחסות גם לתהליכים.

הערכת תוצרים חשובה, מכיוון שהיא מאפשרת לדעת היכן נמצאים חלקים שונים של המערכת, מה יש ומה אין. לא ניתן להימנע מבדיקה כזו. חשוב מאד שתעשה בצורה מקצועית ואמינה, החל מן המורה בכתה. הערכת תוצרים עונה לשאלה "מה". מה קיים, ומה הושג. אם מעוניינים בתשובה של "למה" ושל "איך" עלינו להתבונן בתהליכים ובמורכבויות. בתוך ההערכה של תהליך ניתן למנות את נושא ההערכה החלופית בהערכת הישגים.

יש קשר בין שלושת סוגי המתח האלה, ויש להם נטייה להתכנס. בדרך כלל נראה בעולם שהערכה לפיקוח ובקרה נוטה לבחון בעיקר תוצרים במערכי בדיקה נוקשים. נכון שאפשר גם וגם, נכון שאפשר לשלב. אולם חשוב לזכור שמאחורי כל צד במתחים אלה מסתתרת תפיסת עולם מחקרית-מתודולוגית, אפיסטמולוגית, והרבה פעמים גם ערכית ואידיאולוגית. החיבור הזה של גם וגם הוא לא תמיד פשוט, לא תמיד אפשרי, ובהרבה מקרים פשוט לא נעשה.

לסיכום

מהתבוננות במסמכים הרבים שהוצגו פה ומניתוחם ניתן לראות שאין חשיבה מערכתית ולא מדיניות במה שאמור בידע המקצועי של מורים בתחום ההערכה. לפני פרסום התוכנית הלאומית לחינוך, ההערכה כחלק מתפקידו המקצועי של מורה, ולו במה שאמור בהערכת הישגי תלמידיו, לא נדונה באופן ציבורי. פרסום התוכנית הלאומית יצר שיח בדיבור, בכתובה ובמעשה שהתמקד בסוגיית ההערכה בחינוך. מעט מאד מהשיח הזה התמקד במה צריכים מורים לדעת. מהמסמכים השונים שניתחתי יש רק שניים שעוסקים בידע הנדרש לאנשי חינוך ולמורים: המסמך "הצעה למסגרת לתוכניות לימודים ולפיתוח מקצועי - מדידה והערכה בחינוך" של היוזמה למחקר יישומי בחינוך (הועדה למדידה והערכה בחינוך, 2005), וב"מדריך לתכנון השתלמויות במדידה והערכה בחינוך לשם פיתוח מקצועי של עובדי הוראה" שהוציאה הראמ"ה (הרטף וחוב', 2007). המסמך הראשון הוא בגדר המלצה בלבד, ולשני יכולת השפעה מוגבלת, בעיקר על איכות ההשתלמויות שניתנות בפועל למורים מכהנים, בעיקר במסגרת מרכזי הפסג"ה, אבל לא על עצם קיומן או על כמות ההשתלמויות. עוד פחות יכולים המסמכים להשפיע על פעולות מערכתיות, אקטיביות ורחבות היקף להכשרת עובדי הוראה.

במכללות האקדמאיות להכשרת עובדי הוראה, שהן המכשירות העיקריות של עובדי הוראה בישראל, לא נוסחה נכון לרגע כתיבת המאמר (2009) שום מדיניות להכשרת עובדי הוראה להערכה. מה שנעשה במכללות הוא מקומי וספוראדי ותלוי במקרים רבים במי שנמצא בשטח.

למרות החשיבות הנראית של ידע מקצועי של מורים בהערכה, והיות התחום הזה חלק מהגדרת הידע המקצועי שלהם וזהותם המקצועית, אין היום במדינת ישראל שום מדיניות לגבי הנושא הזה.

מקורות

- אורמיאן, ח. (עורך) (1979). *עמנאל יפה – במעגלי החינוך*. ירושלים: משרד החינוך והתרבות.
- בירנבוים, מ' (1997). *חלופות בהערכת הישגים*. ת"א: רמות
- בירנבוים, מ' (2006). הערכה אלטרנטיבית – לא כל הנוצץ זהב הוא! חינוך בראש העל יסודי, 1, 19-23
- ברזניץ, צ' וימין, ר' (עורכות) (2008) *איבחון מדידה והערכה, גן-ב', תמונת מצב והמלצות*. היוזמה למחקר יישומי בחינוך, ירושלים: האקדמיה הלאומית הישראלית למדעים משרד החינוך והתרבות והספורט יד הנדיב
- בריטברד, ר' וטביביאן-מזרחי, מ' (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 שאוחר ביוני 2009:
- <http://www.academy.ac.il/data/projects/34/E13.pdf>
- הוועדה למדידה ולהערכה בחינוך (2005). *הצעה למסגרת לתוכניות לימודים ולפיתוח מקצועי - מדידה והערכה בחינוך*. היוזמה למחקר יישומי בחינוך, ירושלים: האקדמיה הלאומית הישראלית למדעים משרד החינוך והתרבות והספורט יד הנדיב
- הלוי, א' (2009) *התהוות שדה הערכה בישראל – ההיבט המוסדי*. חיבור כחלק מהדרישות לקבלת התואר "מוסמך למדעי הרוח והחברה" (M.A). אוניברסיטת בן-גוריון נגב: באר-שבע
- ועדת היישום לשילוב הערכה פנימית בבית הספר (2004). *המלצות: במסגרת יישום המלצות כוח המשימה הלאומי לקידום החינוך בישראל*, ירושלים: משרד החינוך, התרבות והספורט.
- הרטף, ח', גנור, מ', רום, א', שלטון, ח' (2007) *מדריך לתכנון השתלמויות במדידה והערכה בחינוך לשם פיתוח מקצועי של עובדי הוראה*. הראשות הארצית למדידה ולהערה (ראמ"ה) משרד החינוך והתרבות.
- התוכנית הלאומית לחינוך (2005). *כוח המשימה הלאומי לקידום החינוך בישראל (דוברת, ש' יו"ר)*. ירושלים: משרד החינוך
- ואן גלדר, א' (2003) *מסמך רקע לדיון בנושא: הכשרת מורים בישראל: מבנה תוכניות 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 8.2-2 (א). ירושלים: משרד החינוך
- חוזר מנכ"ל (2008). *רכז הערכה בית ספרי*. חוזר מנכ"ל 10, סעיף 8.1-1. ירושלים: משרד החינוך
- ויוגב, א', ליבנה, ע', ופניגר, י' (2009) *סקר בינלאומי של הישגים חינוכיים והגלובליזציה של מטרות החינוך*. מגמות, 46 (3)

לוי, א' (ללא תאריך) הלכה ומעשה של תוכניות לימודים. הלכה למעשה, אוסף מאמרים. ירושלים: משרד החינוך, המזכירות הפדגוגית, האגף לתכנון תוכניות לימודים. כפי שאוחזר ב-12 ליוני 2009

http://www.education.gov.il/tochniyot_Limudim/halacha/kategory_5.htm

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

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

לוי-רוזליס, מ' ולפידות, א' (2007) ההכשרה להוראה של מורים מכהנים במרכזי הפסג"ה (תשס"ז). דו"ח מחקר חלק ב', במסגרת רשת עמיתית מחקר. תל-אביב: מכון מופ"ת

לוי-רוזליס, מ' ולפידות, א' (2008) הכשרה להערכה של עובדי הוראה בישראל. דו"ח מחקר מסכם לשלוש שנות מחקר, במסגרת רשת עמיתית מחקר. תל-אביב: מכון מופ"ת

לוי-רוזליס, מ' ולפידות, א' (2009) שדה ההערכה בישראל וזהותם המקצועית של המורים – היש למכללות להכשרת עובדי הוראה סיכוי במשחק? דפים

מירו, א' (2009), "המשותב הבית ספרי" כמנוף לשינויים בית ספריים ומערכתיים. מחקר לשם מילוי חלקי של הדרישות לקבלת תואר "דוקטור לפילוסופיה", הוגש לסנאט אוניברסיטת בן-גוריון בנגב: באר-שבע

נידרלנד, הופמן ודרור (2009). דמות המורה בראי הכשרת המורים: מבט היסטורי. דפים

פרידמן, י' (2005). מדידה והערכה של תוכניות חברתיות וחינוכיות. ירושלים: מכון הנרייטה סאלד.

Alkin, M., (1970) Products for Improving Educational Evaluation, *Evaluation Comment*, vol. 2(3) 1-15

Berger, P. L.; T. Luckmann, (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?* Background paper prepared for the Education for All Global Monitoring Report 2003/4. 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 Evaluation Network and the WMU Evaluation Center. Presentation at the Evaluation Café Presentation at the Evaluation Café, 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. Centre of Excellence for Evaluation, Treasury Board of Canada, Secretariat: Ottawa

Covaleskie, J.F (1994). The educational system and resistance to reform: The limits of policy. *Education Policy Analysis Archives*, 2(4). Retrieved 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 [date] 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
- Goodlad, J., (1995) *Teachers for our Nation's Schools*. Jossey-Bass. San Francisco: CA
- Green, T.F. (1994) Policy questions: A conceptual study. *Education Policy Analysis Archives*, 2(7). Retrieved: <http://epaa.asu.edu/epaa/v2n7.html>.
- Guba, E.G. & Lincoln, Y.S., (1989). *Fourth Generation Evaluation*. Sage Publications. London
- 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*. Rand Corporation: Santa Monica. CA
- Heim, M. (February 1995). "Accountability in Education." (unpublished paper submitted to the Hawai'i 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, pp. 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, 2007, from 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, IN: Phi Delta Kappa Educational Foundation.
- Kellaghan, T., & Greaney, V., (2001) **Using Assessment to Improve the Quality of Education**. Paris: UNESCO.

- 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*. Heinemann: Portsmouth
- Koretz, D., (2003). *Assessing Students With Disabilities: Issues and Evidence*. *Educational Assessment*, 9(1-2) p. 29-60
- 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.
- 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*.
- 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
- 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*. Sage, thousands Oaks. CA
- 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. Pp. 13-103 in Linn, Robert. (Ed.) *Educational Measurement, 3rd Edition*. Washington, DC: American Council on Education.
- 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* (pp. 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.
- 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.
- Stake, R. E.,(1967). The countenance of educational evaluation, *Teachers College Record*, 68,
- 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., (1999). Evaluation culture(s) in Europe: differences and convergence
- Tyler. R. W., (1942). General statement on evaluation. *Journal of Educational Research*. 35. 492-501.