

בשביל מי ובשביל מה? גורמי הטיה בהערכת ההוראה והקורס במכללה לחינוך ושאלת תרומת ההליך למוערכים ולארגון

סיומת אכדור

תקציר

מאמר זה דן בשאלת תפקידו ותוקפו של הליך הערכת ההוראה במוסדות להשכלה גבוהה, ובמיוחד במכללות אקדמיות לחינוך, מנקודות מבט אקדמיות, אתיות ומחקריות. בדיון מוצגים ממצאי מחקר המראים את הבעייתיות בשימוש בלתי מבוקר בממצאי הערכת ההוראה והקורס בשל קיומו של גורם הטיה שיטתי בהערכות, שמקורו בתפיסת המעריך את מאפייניו האפיסטמיים של הקורס המוערך. המחקר עוסק בשלושה היבטים: (א) הערכת מאפייני ההוראה והקורס באמצעות שתי מערכות מדדים וכמה היגדים כלליים על הקורס. (ב) בשאלה האם ניתן לזהות גורמי הטיה שיטתיים בהערכות אלה, ואם כן, מה מקורם. (ג) בשאלה האם רכיבי השאלון מספקים מידע אנליטי ורלוונטי על מאפייני ההוראה והקורס. נמצא כי רכיבי השאלון מאפשרים התרשמות ממוקדת ומפורטת מההוראה ומהקורס. תפיסת המעריכים את אופיו האפיסטמי של הקורס כשייך לאחת משלוש קטגוריות: אקדמי-עיוני-מנותק ממצויאות ההוראה; אקדמי-עיוני-מקושר למציאות ההוראה; מתודולוגי-דידקטי, משפיעה על ציוני ההערכה ועל תפיסת המעריכים את התרומה שהם מפיקים מהקורס. כן נמצא כי דווקא הממד הנמדד יותר בשאלוני הערכה המשמשים בהשכלה הגבוהה, המתייחס ל"איך" מתבצעת ההוראה, אינו גורם המנבא את תפיסת הסטודנטים את תרומות הקורס, בשעה שמאפיינים אפיסטמיים של הקורס הם המנבאים את תרומותיו. ממצאי המחקר והדיון בהיבטים שונים של שימוש בנתוני הערכת ההוראה כהערכה מעצבת וכהערכה מסכמת מתקשרים לשאלות פרקטיות ואתיות בדבר תוקף השימוש בנתונים אלה בהעדר מקורות נוספים להערכה ולשיפוט של איכות ההוראה והקורס.

הערכת ההוראה במוסדות להשכלה גבוהה

מיקוד הדרישה לאחריותיות (accountability) בארגוני חינוך, שמשמעותה אחריות של ממלא תפקיד להליך ולתוצאותיו, עבר מהתמקדות בנושאים כמו: ניהול כספי

וניחול תכניות, להתמקדות בעניינים הנוגעים לאיכות ההוראה ולאיכות מורים (Darling-Hammond, Wise, & Pease, 1983). גם במוסדות להשכלה גבוהה ההתייחסות לאיכות ההוראה היא אחת הדרכים המשמשות להשגת יתר אחריותיות על תהליכי הוראה ועל תפוקות הנגזרות ממנה (Burns, 1998; Huabands & Fosh, 1999; Long, Tricker, Rangecroft, & Gilroy, 1993).

לגורמים שונים במערכת יש עניין שונה בהערכת ההוראה. עניינם של גורמים אדמיניסטרטיביים ואקדמיים, המשתמשים בתוצאות ההערכה כדי לקבל החלטות בנוגע לקביעות, לקידום, לשיפור ההוראה ולהכנת תכניות חדשות, עשוי להיות שונה מעניינם של סטודנטים בהערכה זו, ושונה אף ממה שמעניין את המרצה שהוא מושא ההערכה (Darling-Hammond, Wise, & Pease, 1983; Neumann, 2000). בנוסף לכך קיימים בעלי עניין חיצוניים, כמו גורמי פיקוח אקדמיים ואחרים, גורמים מממנים וגורמים הקולטים את תוצרי המוסדות, המעוניינים בנתונים על איכות המוסד (Burns, 1998; Karapetrovic, Rajamani, & Willborn, 1999; Sirvanci, 1996). בעלי עניין אלה משתמשים במדדים שונים מאלה של גורמים פנים-מוסדיים, אך בהחלט ייתכן כי חלקם יתייחסו גם לאיכות ההוראה במוסד.

איכות ההוראה ומדידתה כמרכיב בהגדרת הפרופסיונליות של סגל ושל מוסדות אקדמיים

הוראה איכותית מקדמת תהליכי למידה ופיתוח כישורים של הלומדים במוסדות להשכלה גבוהה, והאיכויות של בוגרי מוסדות אלה תלויות בין היתר בטיב תהליכי ההוראה שאליהם נחשפו. ספרות הבוחנת את הפרופסיונליות של סגל אקדמי בהשכלה הגבוהה רואה בהוראה עיסוק השייך לסוג העיסוקים הפרופסיונליים המאופיינים בידע מומחיותי. ידע זה מתבטא בין השאר בתוצאות הנמדדות ברמת הידע שרכשו הסטודנטים ובדרכים שבהן הועבר להם ידע זה. העיסוק בהוראה בהשכלה הגבוהה שייך לפרופסיות מסורתיות, כמו רפואה, משפטים ותיאולוגיה, אשר בהן ההוראה, לא פחות מאשר המחקר, נתפסת כמרכיב עיקרי במאפייני הפרופסיונליות ובהעצמתה. כזאת היא מחייבת שליטה והתמחות בתחום ידע דיסציפלינרי, מחקר והתמחות הנוגעים לידע פדגוגי בתהליכי הוראה-למידה של מבוגרים בהשכלה הגבוהה, ביסוס נורמטיבי של התמחויות אלה ושימוש אתי ונבון בידע איך ליישם את הידע (Darling-Hammond, Wise, & Pease, 1983; Farrugia, 1996; Stark & Lattuca, 1997).

במוסדות להשכלה גבוהה בארצות רבות גוברת כיום המגמה לתת משקל רב יותר למרכיב ההוראה בפעילות האקדמית, ומתגבשת התפיסה כי שיפור איכויות ההוראה

דורש מנהיגות בעלת חזון שתשלב ערכים אקדמיים עם שינוי באקלים הארגוני ובתגמולים הניתנים לסגל (Ramsden & Martin, 1996). למעשה, מאז 1990 נדרשים מוסדות להשכלה גבוהה בארצות אלה להוכיח כי הם עומדים בדרישות לאיכות ולביצוע התואמים רמות גבוהות של המושג "אחריותיות". בחינה של איכויות אלה מתבססת על מידת שביעות הרצון של סטודנטים, על נתוני העסקת בוגרים ועל הערכות של תפקוד הסגל (Burns, 1998; Long, Tricker, Rangelcroft, & Gilory, 1999). האחריותיות כלפי הלקוחות השונים מחייבת לא רק תכנון אפקטיבי, שתוצאותיו הן איכות אקדמית בסטנדרטים גבוהים, אלא גם תהליכי בקרה ומדידה של טיב ההוראה. תהליכים אלה משמשים כמכניזם של פיקוח חברתי וכבסיס לבקרה עצמית ולהתפתחות אישית. בקרת איכות על ההוראה כחלק מהגדרת הפרופסיונליות האקדמית של הסגל יכולה לתרום לחיזוק מצוינות פרופסיונלית, להאיר את תרומת הסגל להוראה, למחקר ולחברה, ולהיות בסיס להערכה וליוקרה חברתית (Bayer & Braxton, 1998; Farrugia, 1996; Nevo, 1994). מרצים נדרשים לשמור על יוקרתה של הפרופסיה ועל יוקרת המוסד האקדמי באמצעות התפתחות פרופסיונלית הן בתחומי ידע ייחודי והן בדרכי הנחלתו. זהו אמצעי חשוב להשגת מטרות המוסד מול לקוחותיו ומממנו, והוא מעצים גם את ההוראה האוניברסיטאית כפרופסיה יוקרתית. מוסדות להשכלה גבוהה מפתחים כיום גופים אקדמיים ואדמיניסטרטיביים שמטרתם לעסוק בבקרת איכות ובשיפור ההוראה (Fresko & Nasser, 2001; Neumann, 2000).

הערכת ההוראה מול שאלת החופש האקדמי והזכות לשמירה על צנעת הפרט

למעשה, יש בתהליכי הערכת ההוראה משום סתירה של החופש האקדמי של הסגל במוסדות להשכלה גבוהה, ופרסום תוצאות ההערכה או שימוש בהן עלול להיות נוגד את זכויות השמירה על צנעת הפרט. כך למשל, במוסדות להשכלה גבוהה בארצות הברית ובאוסטרליה יש לקבל אישור מפורש של המרצה כדי שיהיה אפשר להשתמש בתוצאות הערכת ההוראה שלו, וקיימות הגבלות מרחיקות לכת על חשיפת התוצאות לגורמים אחרים מאלה שקיבלו אישור לראותן (Neumann, 2000; Robinson, 1996; Canty, & Fink, 1996). דיני ההשכלה הגבוהה בישראל קובעים כי הגבלת החופש האקדמי עשויה להיות מוצדקת כאשר היא מסייעת להבטיח רמת הוראה נאותה. ברוח זו אושר שימוש בהערכות של סטודנטים באשר לרמת ההוראה של איש הסגל כאמת מידה להחלטות בנושאי קידומו האקדמי. מטרת ההליך היא להימנע מחופש הוראה מוחלט, שעלול לפגוע באיכות ההוראה, ולאפשר החלת מערכת בקרה על האופן

שבו מבוצעות פעילויות הוראה שונות. על פי המדיניות המקובלת במוסדות להשכלה גבוהה, הרשויות נמנעות מלהפעיל אמצעי בקרה ופיקוח חמורים, ולרוב מסתפקות בשאלוני הערכה שהסטודנטים ממלאים (הר-זהב ומדינה, 1999, עמ' 370, סעיפים: 14.33, 14.34). במוסדות אקדמיים להכשרת עובדי הוראה בישראל הגבלת החופש האקדמי היא רבה יותר. על פי הוראת החוק, המכללות האקדמיות להכשרת מורים אינן נהנות מעצמאות ומחופש פעולה אקדמי במידה שניתנת למוסדות אקדמיים אחרים. הן נתונות לפיקוחו של משרד החינוך ושל מפקחים ממונים מטעמו. מעמד ייחודי זה של המכללות האקדמיות לחינוך נובע מכוח הייעוד שלשמו הן קיימות: אספקת עובדי הוראה למערכת החינוך שמקיימת המדינה, ומכוח ייעוד זה גם נדרשת מידה מסוימת של הכתבת תוכני ההוראה (שם, עמ' 171-173, סעיפים: 5.1-5.9).

איך ומה מעריכים?

השימוש בשיטות כמותיות להערכת ההוראה באמצעות שאלוני משוב שמילאו סטודנטים מאפשר עיבוד מהיר של נתונים ושימוש במדדים זהים להערכה מספרית ולהשוואה בין מרצים או להערכת אותו מרצה בזמנים שונים ובקורסים שונים. החסרונות העיקריים של השיטה מתייחסים לכך שההערכה מתמקדת בהיבטים שטחיים משותפים לדיסציפלינות השונות אשר מתבטאים בעיקר בחובות המורה (Scriven, 1995), ולכן היא טכנית ולא מהותית (Pratt, 1997). מרצים מוערכים על פי אותם קריטריונים גם כאשר הקורסים שהם מלמדים שונים מבחינה סכולסטית ואינם ניתנים להשוואה. ממצאי ההערכה יכולים ליצור הכללות מוטות, אשר אינן משקפות את תפקודם האמתי של מרצים ואת כישוריהם (Scriven, 1994). השימוש השכיח בשאלונים אלה מתקיים למרות הביקורת הרבה מצד סגל וסטודנטים כאחד. סיבה עיקרית להתנגדויות הסגל קשורה לבעיות של תוקף הקריטריונים הנכללים בשאלונים ולבעיות של תוקף הכלי. אין הסכמה גם בשאלה מה יש לכלול בהערכת ההוראה ואיך ניתן להגדיר הוראה אפקטיבית (Hutchings, 1996, In: Burns, 1998; Huabands & Fosh, 1993). הניסיון להגיע להגדרות סטנדרטיות בעניין זה אינו בהכרח עולה בקנה אחד עם סוגיית החופש האקדמי של מרצים ועוד פחות מכך עם היבטים אחרים העולים ממטרות ההוראה ומאופי הקורס. יש ספקות באשר למהימנות המידע המתקבל בהליכים אלה, והמחקר מראה כי גורמים שמקורם בציפיות סטודנטים מהקורס, בקשרי מוערכים-מעריכים, במניפולציות הדדיות של סטודנטים על מרצים ולהפך, משפיעים על מהימנות המדידה (למשל: Jacobs, 1987; ו- Marlin, 1987, In: Wachtel, 1998; Sacks, 1996, In: Crumbley, Henry, & Kratchman, 2001).

ביקורת על שיטה זו התפתחה גם בקרב אנשי האסכולה המדגישה שיטות הערכה חלופיות, המעדיפות איסוף מידע אותנטי על תהליכי הוראה-למידה ועל תרומתם ללומדים. אולם גם אסכולה זו זוכה לביקורות וביניהן טענות כי שיטות אלה צורכות זמן רב של איסוף מידע ועיבודו, נוטות להתמקד בעיקר בקשיים, יוצרות חשיפת יתר של סטודנטים להיבטים פחות רלוונטיים המתייחסים להערכת המוסד. בנוסף לכך גם גישות אלה אינן מספקות פתרון הולם לשאלות חשובות, כמו: מהי הוראה אפקטיבית? איך מגדירים אותה מבחינה אופרטיבית? מיהם אותם עמיתים מעריכים? מהי מידת הסובייקטיביות בפרוש הממצאים ומהי תקפותם? בפועל, הן משמשות פחות להערכת ההוראה בהשוואה לשיטה הרווחת של שאלוני המשוב (Hutchings, 1996; In: Burns, 1998; Koon & Murray, 1995).

במכללות האקדמיות לחינוך בישראל משמש המרכיב "מצוינות בהוראה" גורם בעל משקל בהחלטות הנוגעות לקביעות ולקידום, והוא נמדד בין היתר באמצעות נתונים המתקבלים משאלוני משוב על ההוראה. במרבית המקרים הנתונים מתמקדים באיכות הפדגוגית של ההוראה, המשתקפת בהערכות על מבנה הקורס, על קצב ההוראה ועל מידת הבהירות בהצגת הנושאים והמשימות. הספרות מראה כי סטודנטים נוטים להדגיש בעיקר היבטים של סגנון הוראה, יכולת להציג נושא, יכולת להלהיב, הכנה וארגון של השיעור, קיומה של למידה והוגנות במתן ציונים (Crumbley, Henry, & Kratchman, 2001), בשעה שלהיבטים אחרים של ביצוע ההוראה ניתנת מידה פחותה של תשומת לב (Huabands & Fosh, 1993).

ראשיתו של המחקר המוצג במאמר זה הייתה איסוף המלצות מחברי סגל ומסטודנטים במכללה לחינוך על היגדים שרצוי לכלול בשאלון להערכת הוראה. ארבע הקטגוריות העיקריות שהתקבלו מניתוח תוכן של ההיגדים שהציעו חברי הסגל היו: דרכי הוראה (43%), תרומה לפיתוח אקדמי של הלומד (29%), תקשורת עם הסטודנטים (19%) ורלוונטיות הנלמד מבחינה מקצועית (9%). הבולטות הנמוכה יחסית של הקטגוריה האחרונה - "רלוונטיות הנלמד מבחינה מקצועית" - מדגישה את השוני בגישת המרצים מזו של הסטודנטים. משאלון מקביל שהועבר לסטודנטים התקבלו בניתוח תוכן הקטגוריות: דרכי הוראה (48%), רלוונטיות הנלמד להוראה בבית הספר (27%), תקשורת עם הסטודנטים/יחס ללומד (19%), רוחב אופקים של המרצה/רמה גבוהה של מרצים (6%). הממצאים עולים בקנה אחד עם ממצאים נוספים המעידים על שוני בתפיסת הסטודנטים את ההוראה במוסדות להכשרה פרופסיונלית מזו של מוריהם (למשל: שילוח ופרסקו, 1994).

מהימנות ותוקף של הערכת הוראה המבוססת על שאלוני משוב של סטודנטים בשאלוני הערכה סגורים

מדידה והערכה הם שני מושגים המשמשים במאמר זה, וחשוב לעמוד על משמעותם. מדידה (measurement) היא הליך של כימות מידע שבאמצעותו נערך שיוך של סולם ערכים מספרי או של סדרה של קטגוריות לתופעה מוערכת בהתאם לכללים מסוימים. הערכה (assessment) היא תהליך של איסוף מידע הקשור לטיבו של מושא הערכה. טעות הערכה משמעה העדר יציבות בציונים, בדירוגים או בתצפיות. מקור הטעות יכול להיות גורמים שיטתיים, כמו פריטי מבחן לא טובים, תנאי מדידה הפוגמים באיסוף המידע או גורמים לא שיטתיים שאינם ניתנים לפיקוח, למשל: ציפיות המעריכים מתהליך הערכה או ממושא הערכה או מצבם הפסיכולוגי או הפיזי בזמן מתן הערכה. למעשה, בכל הערכה יש מידה מסוימת של טעות (Airasian, 1997). מושגים אלה מתקשרים למושג נוסף - evaluation - שבשפה העברית הוא טבוע במילה הערכה. המושג עוסק בשלב שבו נערך שיפוט לגבי איכותו או טיבו של מושא הערכה, שיפוט המבוסס על נתוני הערכה אך מבטא אמות מידה קונטקסטואליות וערכיות (Aspinwall, Simkins, Wilkinson, & McAuley, 1992), המשמשות בעת קבלת החלטות הנוגעות למושא הערכה.

הערכת ההוראה בחינוך הגבוה, המבוססת על שאלוני משוב שממלאים סטודנטים, היא התחום הנחקר ביותר בנושא הערכת ההוראה (Cashin, 1988, In: Wachtel, 1998). מהימנות מתייחסת למידת העקביות של המדידות ומוצגת בערכים מתמטיים. השימוש הרווח בשאלונים סגורים להערכת הוראה נשען על מחקרים המראים כי סטודנטים בהשכלה הגבוהה מגיבים לטיבה של ההוראה. יש להם תפיסה טובה בנוגע לטבעה ולבהירותה, וזו נמצאה יציבה מעבר לזמן ומעבר לקורסים (Hativa & Raviv, 1998, In: Hativa, 1998). כלומר במדידות חוזרות לאורך זמן ובתנאים דומים רמת הציונים נשארת דומה. עם זאת, השיטה מעמידה שאלות של תקפות שיש להתמודד אתן (Barrett, 1986; Dunkin & Barnes, 1986; Darling-Hammond, Wise, & Pease, 1983; Pratt, 1997). כפי שיפורט בהמשך, תקפות של כלי המדידה היא המידה שבה הכלי מודד את שהוא אמור למדוד מבחינת תוכן. המשתנים המשפיעים במיוחד על הערכת ההוראה קשורים לציפיות הסטודנטים מההוראה לאור צורכיהם הלימודיים, לתפיסת מרצים את תפקידם בהוראה וביטוייה בפועל ולאיינטראקציה בין שתי קבוצות משתנים אלה. מהימנות נחשבת לחלק מהדרישה לתקפות. כדי שכלי יהיה תקף עליו להיות עקבי במדידה, אך הוא יכול להיות מהימן מבלי שיהיה תקף (נבו, 2001). הטיה שיטתית במדידה, שניתן לזהותה, פוגעת בתוקפו של כלי המדידה.

גורמים העלולים לפגום במהימנות ובתקפות של הציונים המתקבלים משאלוני משוב להערכת הוראה

הממצאים הבולטים בספרות המחקרית, הנוגעים לבעיות של הטיה ופגימה במהימנות ובתוקף של ציוני ההערכה המתקבלים משאלוני משוב להערכת ההוראה, קובצו בסקירות ספרות מפורטות שפרסמו בעשור האחרון האסבנדס ופוש (Husbands & Fosh, 1993) ווצ'טל (Wachtel, 1998). הרשימה מציינת הטיות בציוני ההערכה הנובעות מתנאי המדידה והטיות הנובעות מגורמים המשפיעים על ההערכה. ברשימת הגורמים הבולטים המשפיעים על הטיה במדידה הנמצאת ביסוד ההערכה, בכיוון של העלאת גובה הציונים, תוארו המצבים האלה: העברת השאלון בסמוך למועד הבחינה בקורס; העדר אנונימיות של המעריכים; נוכחות המרצה בכיתה או החשש כי הוא יראה את ההערכות; כיתות מאוכלסות במספר מועט של תלמידים. רשימת הגורמים המשפיעים על הטיה בהערכה בכיוון של ירידה בגובה הציונים כוללת: מאפיינים אפיסטמיים של הקורס; קורסי חובה (בניגוד לקורסי בחירה); קורסים במתמטיקה ובמדעים וקורסים "קשים" ללמידה (בהשוואה ל"קלים" יותר). בין הגורמים הבולטים בהטיית הציונים שמקורם במעריכים נכלל גורם המתייחס לתפיסת הסטודנטים את אישיות המרצה בקורס. בעניין זה חוקרים מציגים היבט מעניין בשאלת התקפות של כלי המדידה. לשיטתם, רגישות להיבט האישי מועלה את התקפות של הכלי, ולכן ממצאי הערכה בקורסים שונים של אותו מרצה מספקים תמונת תפקוד תקפה יותר. כן נמצא שסטודנטים "באים בחשבון" עם מרצים שאינם עונים על ציפיותיהם (Crumbley, Henry, & Kratchman, 2001), וכי התעניינות מוקדמת של הסטודנטים בנושאי הקורס ומוניטין טוב של מרצים משפיעים לטובה על ההערכות (Babad, Darley, & Kaplowitz, 1999; Wachtel, 1998). בעניין זה ראוי לציין כי במוסדות להכשרה פרופסיונלית יש לסטודנטים ציפיות ברורות למדי מתכנית הלימודים, והם מעריכים אותה בעיקר על פי מידת היותה רלוונטית לצורכיהם (אבדור, 2000; Stark & Lattuca, 1997; Scriven, 1994).

מדוע בכל זאת אפשר להשתמש בשאלון אחיד מטיפוס כמותי לאיסוף מידע על איכות ההוראה מנקודת מבטם של הסטודנטים? חלק נכבד מהחוקרים הדנים בבעיות של תוקף השאלון וממצאיו בכל זאת אינם ממליצים לבטל לחלוטין את השימוש בו. בין הסיבות לכך בולטת הטענה, שההערכות יכולות לשמש מקור מידע אחד לצד מקורות נוספים המהווים יחד מאגר מידע על אודות מאפיינים ואיכויות של קורסים, של תכניות לימודים ושל תהליכי הוראה. לפי עמדה זו, תהליכי הערכת ההוראה הנדרשים לצורך בחינת ממד האחראיות מחייבים שימוש במידע סטנדרטי, אובייקטיבי וחיצוני,

שמקורו אינו במורה עצמו. איסוף מידע זה צריך שיעשה במדרג ארגוני היררכי ומבוקר כדי להבטיח אמינות ואחידות. יש להתחשב בהקשר שבו מתבצעת הפעילות המוערכת ולהבטיח שהמידע אכן מתאים ומספק, וייתכנו מקרים שבהם צריך להשתמש בשיטות הערכה מגוונות (Darling-Hammond, Wise, & Pease, 1983). כאשר קיימת מערכת נוספת לבקרת איכות אקדמית של תהליכי ההוראה הבודקת את איכות תכנית הקורס, את חומרי הלמידה, את הדרישות מהסטודנטים ואת קיומם של תנאים הולמים להוראה (גודל כיתה, עזרים), מהווה הערכת ההוראה רק זווית אחת, מנקודת מבטו של הסטודנט כצרכן שירות, מתוך כמה זוויות שדרכן נבחנת איכותו של המוסד האקדמי (Koon & Murray, 1995).

המחקר

מטרת המחקר ושאלות המחקר

הבעייתיות שתוארה בסקירת הספרות בדבר תוקף שאלוני הערכת ההוראה והקורס מטרידה גם את מושאי ההערכה וגם את מקבלי ההחלטות הנסמכים לא מעט על ממצאים אלה. נמצא כי סטודנטים בתכניות להכשרה להוראה באוניברסיטאות ובמכללות האקדמיות לחינוך מעדיפים בלימודי החינוך קורסים שאופיים מרשמי. תפיסה זו "צובעת" את החוויה הקוריקולרית שהם מתארים באשר למאפיינים אפיסטמיים ואחרים של הקורסים (אבדור, 2001). בזיקה לממצאים אלה, הנחת המוצא של המחקר הנוכחי היא שקיימת הטיה בהערכת ההוראה באמצעות שימוש בשאלוני משוב של סטודנטים, המבוססים על הערכה כמותית, שכן הערכה זו מושפעת בין היתר ממידת שביעות רצונם של המעריכים ממאפיינים אפיסטמיים של הקורס. למעשה, ביסוד תהליך הערכת ההוראה בשיטה זו קיימת השפעה סמויה של משתנים הקשורים שמקורם במאפייני ההוראה, הקורס והמעריכים (Kember, 1997; Stringer & Irwing, 1998).

מטרת המחקר הייתה לפתח שאלון להערכת מאפייני ההוראה והקורס בשני ממדים: איך הקורס מתוכנן ומועבר על ידי המרצה (התחום העיקרי המאפיין את מרבית השאלונים מטיפוס זה), ומה אופיו האפיסטמי מבחינת בסיסי הידע המשמשים בו: ידע שהוא בעיקר עיוני-אקדמי, ידע שהוא בעיקר פרקטי או שילוב בין שני סוגי הידע. פיתוח השאלון התבסס על שלוש שאלות מחקר שמטרתן הייתה להעמיק ולחקור את סוג המידע המתקבל משאלוני הערכת ההוראה כדי לדעת כיצד להתייחס למהימנותו ולתוקפו. הראשונה, האם סטודנטים הלומדים באותו קורס ואצל אותו

מרצה עשויים לתפוס את מאפייניו האפיסטמיים באופן שונה? אם התשובה לכך היא, כמצופה, חיובית, נשאלת השאלה השנייה: האם הבדלים בתפיסת הסטודנטים את אופיו האפיסטמי של הקורס מתבטאים גם בהבדלים בציוני ההערכה שהם נותנים להוראה ולקורס? השאלה השלישית הייתה: האם ההבחנה בשאלון ההערכה בין ממד ה"איך" לממד ה"מה" וכן היבטים נוספים שנכללו בשאלון זה מספקים מידע רלוונטי ומובחן על המרצה, על הקורס ועל תפיסת הסטודנטים את תרומות ההוראה והקורס להכשרה להוראה ולהתפתחות האישיה?

חשוב להדגיש כי פיתוח שאלונים להערכת ההוראה צריך שיותאם לאופיים של הקורסים המוערכים, למטרותיהם ולמאפייני ההוראה/הנחיה בהם. השאלון שפותח במחקר זה מתאים להערכת ההוראה בקורסים עיוניים וסדנאיים מטיפוסים שונים, הנלמדים בקמפוס האקדמי. אין הוא מתאים להערכת ההוראה בסמינריונים, בלימודים מקוונים ובהדרכה המעשית לסוגיה.

השיטה

הבחנה בין סוגי קורסים באמצעות תפיסת סטודנטים את מאפייניהם האפיסטמיים

בסקירת הספרות הוצגו גורמי הטיה בהערכת קורס שמקורם בהשפעה ישירה של משתנים הקשורים במאפייני ההוראה, באופי תוכני הקורס, במוטיבציה של סטודנטים להתמודד עם התכנים, בצורכיהם הסוציאליזטוריים של הסטודנטים ובהשפעות גומלין בין מערכות משתנים אלה. תכניות להכשרת מורים כוללות קורסים בעלי אופי אפיסטמי שונה, החל בקורסים שאופיים עיוני-מופשט ועד לקורסים שאופיים פרקטי במובהק. בהערכת ההוראה יש צורך להבחין בין סוגי קורסים אלה. קריטריון מבחין ורלוונטי נמצא בהבחנה בין קורסים "מכווני קמפוס", שאופיים עיוני ומנותק מהפרקטיקה, לבין קורסים "מכווני שדה" שאופיים פרקטי בעיקרו (Zeichner & Gore, 1990). את תקפות ההבחנה בין סוגי קורסים מאששים ממצאים המראים כי סטודנטים הלומדים בתכניות להכשרה פרופסיונליות נוטים לתת למרצים הערכות גבוהות יותר מאלה הניתנות בקורסים אוניברסיטאיים, שכן הם עונים על צורכיהם הסוציאליזטוריים (Scriven, 1994). כמו כן, בתכניות להכשרת מורים השפעת אופיו האפיסטמי של הקורס על הערכתו רבה יותר מאשר בקורסים דומים הניתנים במסגרת אוניברסיטאית (אבדור, 2000). במחקר המדווח במאמר זה נעשתה הבחנה בין סוגי קורסים על פי התייחסות המעריכים למאפייניהם האפיסטמיים באמצעות סיווגם את הקורס המוערך לאחת משלוש קטגוריות אפשריות: קורס עיוני המנותק

ממציאות ההוראה (עמנ); קורס עיוני המקושר למציאות ההוראה (עמק); קורס מתודולוגי-דידקטי (מתד) (אבדור, 2001).

מבנה השאלון ואופן מילוי ההערכה

בשאלון 20 היגדים שזוהו בניתוח גורמים כמתייחסים לממדי ה"איך" (הקורס מנוהל על ידי המרצה) וה"מה" (אופיו האפיסטמי) המשמשים בהוראה. הגורם הראשון שייך לבסיס הידע האקדמי בקורס (ה"מה") ומסביר 61.4% מהשונות. הגורם השני קשור לניהול הקורס (ה"איך") ומוסיף עוד 8.1% לשונות. הגורם השלישי מתייחס לבסיס הידע הפרקטי בקורס (ה"מה") ומוסיף 3.8% לשונות. שלושת הגורמים מסבירים 73.3% מהשונות.¹ ההיגדים בכל גורם הם:

1. **ניהול הקורס** (9 פריטים; אלפא קרונבך = 0.93):

- ◀ המרצה מגיב לשאלות הסטודנטים;²
 - ◀ המרצה מתייחס בכבוד לדעות השונות מדעתו;
 - ◀ למרצה ידע רב בתחום הנלמד;
 - ◀ המרצה משמש מודל למצוינות בהוראה;
 - ◀ המרצה קושר בין הנושאים הנלמדים לבין תחומים אחרים;
 - ◀ המרצה ממחיש את התכנים הנלמדים בעזרת דוגמאות;
 - ◀ הנושאים הנלמדים מוצגים באופן מסודר ומובנה;
 - ◀ הדרישות הלימודיות מוגדרות בבירור;
 - ◀ המטלות בקורס (קריאה, תרגיל או עבודה) נותנות ביטוי לידע שנרכש בו.
2. **מאפייני ידע עיוני-אקדמי המשמש בקורס** (6 פריטים; אלפא קרונבך = 0.93):
- ◀ המרצה ממחיש את הנלמד באמצעות מחקרים או פרסומים אקדמיים עדכניים;
 - ◀ הקורס תרם ליכולתי לבצע מטלות אקדמיות שרמתן גבוהה;
 - ◀ בקורס הודגם הקשר בין חשיבה תאורטית לבין חקירה מדעית;
 - ◀ הקורס תרם ליכולתי לנסח שאלות לעיון ו/או למחקר בנושאים שהוצגו בו;
 - ◀ הקורס גרם לי להתעניין בנושאי, מעבר למחויבותי לבצע מטלות שניתנו בו;
 - ◀ הקורס שיפר את מיומנותי לקרוא חומר עיוני ברמה גבוהה.
3. **מאפייני ידע פרקטי המשמש בקורס** (5 פריטים; אלפא קרונבך = 0.95):
- ◀ הקורס שיפר את יכולתי להתבונן באופן רפלקטיבי על עבודתי בהוראה;

1. לפירוט פיתוח השאלון, ובכלל זה ממצאי ניתוח של גורמים שהיו בסיס לגיבושו הסופי, ראו: אבדור, 2005.
2. ההיגדים מנוסחים בלשון זכר, ומכוונים למרצים ולמרצות כאחד.

- ◀ ניתן להשתמש בידע שנרכש בקורס לתכנון דרכי עבודה בהוראה;
- ◀ התכנים הנלמדים מקושרים למציאות העבודה בתחום החינוך;
- ◀ הקורס עזר לי להבין טוב יותר את מאפייני תפקודי בהוראה ובחינוך;
- ◀ בקורס נלמדו דרכים לביצוע מטלות שונות של תפקיד המורה.

בנוסף לכך השאלון כולל שישה היגדים כלליים: המלצה על הקורס, נכונות להשתתף בקורסים נוספים בהנחיית אותו מרצה, תרומת הקורס להתפתחות האישית, תרומת הקורס להתפתחות המקצועית, תרומת הקורס להרחבת ידע בנושאים שנלמדו בו, תרומת הקורס להרחבת ידע בנושאים שמעבר לנושאים שנלמדו בו.

סולם ההערכה הוא בן שש דרגות, ורק קצותיו מוגדרים מילולית. הציון "6" מבטא את הרמה הגבוהה ביותר בהערכה ("במידה רבה מאוד"), והציון "1" מבטא את רמת ההערכה הנמוכה ביותר ("בכלל לא"). לאחר שלב הערכת הקורס התבקשו המעריכים לסווג את הקורס המוערך לאחת משלוש הקטגוריות שצוינו לעיל: קורס עיוני המנותק ממציאות ההוראה (עמנ); קורס עיוני המקושר למציאות ההוראה (עמק); קורס מתודולוגי-דידקטי (מתד).

לוח 1: מתאמים בין שלושת מדדי ה"איך" וה"מה"

מאפייני ידע פרקטי ("מה")	ניהול הקורס ("איך")	
	-	ניהול הקורס ("איך")
	.74	מאפייני ידע פרקטי
.86	.76	מאפייני ידע עיוניים-אקדמיים

המתאמים בין מדדי ה"איך" למדדי ה"מה" נמצאו גבוהים למדי. טווח ערכיהם נע בין .74 ל-.86. (לוח 1). למרות גובה המתאמים בין המדדים אין בממצאים אלה כדי לבטל את תוקף ההבחנה ביניהם, שכן כפי שיוצג בהמשך, הם שני גורמים התורמים לסטודנט אך שונים זה מזה מבחינת תפיסתו את הקורס כתורם להכשרתו להוראה, לעומת תפיסתו את הקורס כתורם להתפתחותו האישית. הממצאים בדבר מתאמים גבוהים בין מדדים שבאמצעותם מוערכת ההוראה בשאלונים מסוג זה עולים בקנה אחד עם מחקרים המתארים את ההוראה כמורכבת מכמה תפיסות המתייחסות למהותה. תפיסות אלה אינן מובחנות זו מזו בבירור, ומרצים מחזיקים בהן בו-זמנית תוך שהן משפיעות על אופן ביצועם את ההוראה (Kember, 1997; Murray & McDonald, 1997). מחקרם של סטרינג'ר ואירווינג (Stringer & Irwing, 1998)

מציג מתאמים גבוהים מאוד בין גורמים/מדדים בשאלון הבודקים היבטים שונים בהערכת ההוראה במוסדות להשכלה גבוהה, אשר הגיעו גם לערך 92. החוקרים דנים בגובה המתאמים ומציינים כי הגורמים הוגדרו היטב מבחינה תוכנית, ולמרות היותם קשורים זה בזה סטטיסטית (כפי שהדבר מתבטא בגובה המתאמים ביניהם) הם מבחינים בין היבטים שונים של תהליך ההוראה ומסייעים רבות להערכתה.

המדגם

המדגם מבוסס על הערכות הוראה-קורס של 180 קורסים במהלך שנת הלימודים תשס"ב (2002). מספר המעריכים היה 3,107, מתוכם 2716 (87%) סיווגו את הקורס המוערך לאחת משלוש הקטגוריות: **עמנ**, **עמק** או **מתד**. הקורסים במדגם הוערכו בכפוף להחלטת ההנהלה שנבעה מאחת או יותר מהסיבות האלה: צורך לדון במועמדות המרצה לקידום או לקבלת קביעות, המרצה חדש במערכת, המרצה לימד קורס חדש.³ כמו כן הוערכו קורסים שהמרצים עצמם ביקשו בהם הערכה כדי להשתמש בתוצאותיה לצורכיהם. במדגם המחקר לא נכללו קורסים שהוגדרו בקוריקולום כקורסים בדיקטיקה של מקצועות ההוראה,⁴ כדי לבחון את אופן תפיסת הקורס בעיני הסטודנטים כפי שהיא מתבטאת בסיווג לאחת משלוש הקטגוריות (**עמנ**, **עמק** או **מתד**).

ממצאים

א. סיווג הסטודנטים את הקורס לאחד משלושת הסוגים: **עמנ**, **עמק** או **מתד**

סיווג הקורסים העיוניים לאחד משלושת סוגי הקורסים, **עמנ**, **עמק** או **מתד**, נבחן בשתי דרכים: (א) נבדקה התפלגות סיווג הקורסים לשלושת הסוגים בכל דיסציפלינה/תחום (לוח 2). הממצאים מראים מגמה בולטת בסיווג: מרבית הסטודנטים סיווגו את הקורסים לקטגוריית הקורסים מטיפוס **עמק** (כ-67% מההערכות). קיימת זהות בשיעורי המעריכים שסיווגו את הקורסים לטיפוס **עמנ** ולטיפוס **מתד** (16%-17%). שתי קבוצות של סטודנטים משתי התמחויות: חינוך מיוחד ומינהל החינוך, בלטו בסיווגם את הקורסים כשייכים לטיפוס **עמק**: 88% ו-74% מההערכות, בהתאמה.

3. המחקר נערך רק על הקורסים הנלמדים במגזר היהודי מסיבות שמקורן בבעיות בהבנת השאלון בעברית במגזר הבדואי. תרגום השאלון לערבית לא היה מספק וחשף בעיות בהבנת משמעויות ההיגדים. כמו כן נצפו בהערכות הסטודנטים במגזר זה דפוסי התייחסות אחרים לקורס ולמרצה, שמקורם בגורמים נוספים על אלה שפורטו בסקירת הספרות.
4. הוצאת קורסים אלה מהמדגם אפשרה למנוע "זיהום" (קונטמינציה) בבדיקת סיווג הקורס המוערך לאחת משלוש הקטגוריות ולבחון את הקשר בין תפיסת אופיו האפיסטמי של הקורס לבין מגמות בהערכתו.

(ב) עבור כל קורס חושב מדד קונצנזוס המבטא את מידת ההסכמה של סטודנטים שלמדו בו על מאפייניו האפיסטמיים. קורס סווג כשייך לקבוצה שבה יש קונצנזוס גבוה בנוגע למאפייניו האפיסטמיים, אם לפחות 75% מהמעריכים סיווגו אותו לאותו סוג של קורסים. הממצאים מראים כי קיים פיזור גדול (בטווח שבין 35% ל-92%) במידת הקונצנזוס באשר למאפייניהם האפיסטמיים של הקורסים בדיסציפלינות/ בתחומים השונים. גם בעניין זה בלימודי החינוך המיוחד יש קונצנזוס כמעט מלא באשר לאופיים של הקורסים (ב-92% מהקורסים לפחות 75% מהמעריכים סיווגו את הקורס באופן דומה). מדדי הקונצנזוס נמוכים הרבה יותר בתחומים האחרים, תופעה המעידה על שונות גדולה בין הסטודנטים הלומדים בהם בתפיסתם את מאפייניהם האפיסטמיים.

לוח 2: התפלגות סיווג הקורסים לאחד משלושת הסוגים ונתוני "מדד הקונצנזוס" בכל דיסציפלינה/תחום

ממד הקונצנזוס	התפלגות סיווג הקורסים על ידי הסטודנטים (באחוזים)			מספר הקורסים	דיסציפלינה/תחום
	עמנ (1)	עמק (2)	מתד (3)		
58	17	64	19	65	לימודי חינוך (אינם מוגדרים כדידקטיקה)
35	10	66	24	17	פדגוגיה
92	2	88	9	13	חינוך מיוחד
55	10	74	16	11	מינהל החינוך
38	21	63	16	26	הומניסטיקה
39	25	62	13	18	שפות זרות (אנגלית, ערבית)
40	17	66	17	30	מדעים (מתמטיקה ומדעי הטבע)
-	16%	67%	17%	180	סך הכול

ב. הבדלים בהערכות ההוראה ובמאפייני הקורס בזיקה לסיווג מאפיינים אפיסטמיים של הקורס

נמצא קשר בין גובה ההערכות בקורסים בשלושת המדדים ובששת ההיגדים הכלליים לבין סיווג הקורס לאחד משלושת סוגי הקורסים.

12. הבדלים בממוצעי ההערכות במדדי ה"איך" וה"מה"

בוצע ניתוח שונות חד-כיווני (ANOVA) עבור כל אחד משלושת המדדים לפי סיווג

הקורס לאחד משלושת סוגי הקורסים.⁵ בכל הניתוחים התקבלו הבדלים ברמה של $p < .0001$ (לוח 3).

לוח 3: ממוצעים וסטיות תקן של מדדי ה"איך" וה"מה"

P	מתד N = 458		עמק N = 1818		עמנ N = 440		אלפא קרויבך	מדדים
	ממוצע	ס"ת	ממוצע	ס"ת	ממוצע	ס"ת		
(**) < .0001	0.9	5.31	0.8	5.37	1.2	4.49	.93	ניהול הקורס ("איך")
(**) < .0001	1.3	4.75	1.2	4.87	1.4	2.88	.95	רלוונטיות להוראה ולבית הספר ("מה")
(*) < .0001	1.4	4.45	1.3	4.63	1.5	3.22	.93	מאפיינים עיוניים-אקדמיים ("מה")

(*) במבחן TUKEY: שלושת הממוצעים שונים זה מזה.
 (***) במבחן TUKEY: ממוצעי הקורסים מהטיפוסים עמק ומתד שונים מממוצע הקורסים מטיפוס עמנ.

מהשוואות אלה עולה כי סטודנטים שסיווגו את הקורס המוערך כשייך לקורסים מטיפוס **עמנ**, העריכו את הקורס בכל המדדים בהערכות נמוכות באופן מובהק בהשוואה להערכות שהתקבלו כאשר הקורס סווג לשני טיפוסים האחרים. באופן מפורט יותר הממצאים מראים כי:

- ◀ בממד "ניהול הקורס" ("איך") בקורסים שסיווגו המעריכים כשייכים לטיפוס **עמנ**: ממוצע הציונים נמוך באופן מובהק ($p < .0001$) מממוצע הציונים שניתנו לקורסים שסווגו לשני טיפוסים האחרים, שהם קורסים המקושרים למציאות ההוראה (ציון 4.49 לעומת הציונים 5.37 ו-5.31, בהתאמה).
- ◀ גם בממד "מאפייני ידע פרקטי" ("מה") התקבלה אותה מגמה: בקורסים שסיווגו המעריכים כשייכים לטיפוס **עמנ** - ממוצע הציונים נמוך באופן מובהק ($p < .0001$) מממוצע הציונים שניתנו לקורסים שסווגו לשני טיפוסים האחרים (ציון 2.88 לעומת הציונים 4.87 ו-4.75, בהתאמה).
- ◀ בממד "מאפיינים עיוניים-אקדמיים" ("מה") נמצאו הבדלים מובהקים ($p < .0001$) בגובה ההערכות בין שלושת סוגי הקורסים: הציונים הגבוהים נתקבלו בהערכות שבהן סווג הקורס לטיפוס **עמק** (ציון 4.63); בתוכם - ההערכות שסיווגו את הקורס כשייך לטיפוס **מתד** (ציון 4.45); הציונים הנמוכים ביותר נתקבלו כאשר הקורס סווג כשייך לטיפוס **עמנ** (ציון 3.22).

5. הבחירה בפרוצדורה הסטטיסטית ANOVA ולא ב-MANOVA נבעה מהצורך לבחון האם קיימים הבדלים מובהקים בין שלושת הממוצעים שחושבו עבור כל אחד משלושת סוגי הקורסים, ולזהות את הממוצעים החריגים בפרוצדורת TUKEY (ברמת מובהקות של 0.05).

ממצאים בעייתיים אלה, ובמיוחד הממצא הפרדוקסלי הנוגע להערכות של מידת השימוש במאפייני ידע עיוניים-אקדמיים בקורסים שסוגו כשייכים לטיפוס **עמנ**, מחייבים הבהרה. דווקא בקורסים הנתפסים בעיני הסטודנטים כמאופיינים בידע עיוני-אקדמי - ממוצע הציונים במדד המתייחס למאפייני ידע זה הוא נמוך (3.22)! ממצא זה מצטרף לממצאים דומים בספרות המחקרית שצוטטה לעיל, ומחזק ממצאים ספציפיים בדבר ההטיה הטבעה בהערכת ההוראה בכלל, ובהכשרה פרופסיונלית בפרט. הטיה זו קשורה כנראה למידה נמוכה של שביעות רצון מהקורס, הנובעת או מאופיו האפיסטמי או מהעדר נכונות, או אפילו יכולת אקדמית, להתמודד עם תכניו, או משני היבטים אלה גם יחד.

2.2. הבדלים בממוצעי ההערכות בהיגדים הכלליים

גם בששת ההיגדים הכלליים נמצאו הבדלים בין סוגי הקורסים בזיקה לאופן סיווג הקורס על ידי המעריכים ובאותו דפוס שהתקבל בשלושת המדדים (לוח 4).

לוח 4: ממוצעים וסטיות תקן של הערכת היבטים כלליים בקורס

P	מתד N = 458		עמק N = 1818		עמנ N = 440		היבטים כלליים
	ממוצע	ס"ת	ממוצע	ס"ת	ממוצע	ס"ת	
(< .0001 *)	1.5	4.87	1.2	5.16	1.8	3.75	תרומת הקורס להרחבת ידע בנושאי
(< .0001 *)	1.6	4.63	1.4	4.89	1.8	3.5	תרומת הקורס לידע מעבר לנושאי הספציפיים
(< .0001 *)	1.6	4.64	1.4	4.83	1.6	2.65	תרומת הקורס להכשרת ההוראה
(< .0001 *)	1.6	4.62	1.4	4.89	1.9	3.36	תרומת הקורס להתפתחות האישית
(< .0001 (**))	1.6	4.85	1.5	5.01	2.0	3.51	נכונות ללמוד בקורסים אחרים של המרצה
(< .0001 (**))	1.6	4.89	1.5	5.04	2.0	3.65	המלצה לאחרים להשתתף בקורס

(*) במבחן TUKEY: ממוצעי שלושת הקורסים שונים זה מזה.
 (***) במבחן TUKEY: ממוצעי הקורסים מהטיפוסים עמק ומתד שונים מממוצע הקורסים מטיפוס עמנ.

בכולם נמצא כי ציוני ההערכה הנמוכים ביותר ניתנו לקורסים שסוגו לטיפוס **עמנ** ($p < .0001$). ניתן לראות כי:

◀ בשני ההיגדים הבודקים את תפיסת המעריכים את תרומת הקורס להתפתחותם (האישית והמקצועית) ובשני ההיגדים הבודקים את תפיסת הקורסים כמעשירים

בידע (בתוכני הקורס ובנושאים אחרים) - ההערכות שונות במובהק זו מזו בכל שלושת סוגי הקורסים ($p < .0001$): סטודנטים שסיווגו את הקורס כשייך לטיפוס **עמק** העריכו אותו בציוני ההערכה הגבוהים ביותר, אחריו הוערכו קורסים שסווגו לטיפוס **מתד**, והציונים הנמוכים ביותר ניתנו לקורסים שסווגו כשייכים לטיפוס **עמנ**.

◀ בשני ההיגדים הבודקים המלצות ללמוד אצל המרצה (באותו קורס או בקורסים אחרים) - הציונים בקורסים שסווגו כשייכים לטיפוס **עמנ** נמוכים במובהק מהציונים שניתנו לקורסים שסווגו לאחד משני הסוגים האחרים ($p < .0001$).

3.2. תרומת המדדים וההיגדים הכלליים לתפיסת המעריכים את הקורס כתורם להתפתחות האישית ולהכשרה להוראה

לוח 5: "תרומת הקורס להתפתחותי האישית" ו"תרומת הקורס להכשרתי להוראה" - משקלי ביתא ואחוז שונות מוסברת בשיטת רגרסיה בצעדים על המשתנים התלויים

תרומת הקורס להכשרתי להוראה				תרומת הקורס להתפתחותי האישית			
P	RSq	β		P	RSq	β	משתנה
<.0001	.766	.674	1. מדד: מאפייני ידע פרקטי ("מה")	<.0001	.674	.324	1. היגד: הקורס תרם להרחבת הידע שלי מעבר לתחום עיסוקו
<.0001	.790	.223	2. היגד: הקורס תרם להרחבת הידע שלי בתחום שבו הוא עוסק	<.0001	.743	.188	2. מדד: מאפייני ידע עיוניים-אקדמיים ("מה")
<.03	.790	.040	3. מדד: מאפייני ידע עיוניים-אקדמיים ("מה")	<.0001	.763	.204	3. היגד: הקורס תרם להרחבת הידע שלי בתחום שבו הוא עוסק
				<.0001	.770	.156	4. מדד: מאפייני ידע פרקטי ("מה")
				<.0001	.773	.053	5. היגד: הייתי מעוניין ללמוד עם מרצה זה גם בקורסים אחרים
				<.0001	.774	.047	6. היגד: הייתי ממליץ ללמוד את הקורס אצל מרצה זה

נערכו שני ניתוחי רגרסיה מטיפוס "רגרסיה בשלבים" על שני משתני קריטריון שנכללו בשאלון. מטרת הניתוחים הייתה לבדוק תפיסה כללית ומסכמת של

6. רגרסיה בשלבים (stepwise regression) - ללא הכתבה של סדר כניסת המשתנים למשוואה. כמו כן נבדקו הנתונים ברגרסיה סטנדרטית. השוואת הממצאים של שני סוגי הרגרסיות מראה, כי לא היה ביניהן הבדל מבחינת אחוז השונות המוסברת (RSq) ולא מבחינת המשתנים העיקריים התורמים להסבר המשתנה התלוי. הבחירה ברגרסיה בצעדים מאפשרת זיהוי של התרומה לשונות המוסברת של כל משתנה הנוסף למשוואה, ובכך גם להשוות את המשתנים המנבאים בשתי הרגרסיות ואת תרומתם למשוואת הניבוי.

הסטודנטים את הקורס מבחינת תרומתו להתפתחותם האישית ולהכשרתם להוראה (לוח 5). המשתנים הבלתי תלויים במשוואות הרגרסיה הם: מדד ניהול הקורס ("ה"איד"), שני מדדי מאפייני הידע ("ה"מה"), שני היגדים העוסקים בהמלצה על הקורס בהנחיית המרצה ושני היגדים הקשורים לתרומת הקורס להרחבת הידע בתכניו הספציפיים ולתרומתו להרחבת ידע שמעבר לתכניו.

ממצא עקרוני ומעניין במיוחד שהתקבל בשתי הרגרסיות הוא כי המדד העוסק ב"ניהול הקורס" ("איד"), שהפריטים הנכללים בו הם השכיחים ביותר בשאלונים להערכת ההוראה, כלל לא נכנס למשוואות הרגרסיה כמשתנה מנבא. ממצאים נוספים הם:

◀ אחוז השונות המוסברת של המשתנים התלויים על ידי המנבאים גבוה בכל אחת מהרגרסיות, ויש הבדל במאפייני המשתנים המנבאים כל אחד מהמשתנים התלויים.

◀ ברגרסיה על "תרומת הקורס להתפתחותי האישית" אחוז השונות המוסברת הוא 77.4%. בולטים בה מקדמי ביתא של שלושת המשתנים המנבאים (לפי סדר כניסתם לרגרסיה) שאופיים אקדמי-סכולסטי: ההיגד "הקורס תרם להרחבת הידע שלי מעבר לתחום שבו הוא עוסק" ($\beta = .324$); המדד "מאפייני ידע עיוניים-אקדמיים" של הנלמד בקורס ($\beta = .188$); וההיגד "תרומת הקורס להרחבת הידע בתחום שבו הוא עוסק" ($\beta = .204$). שלושה משתנים אלה מסבירים 76.3% מהשונות. התוספת של עוד שלושה מנבאים: המדד "מאפייני ידע פרקטי" וההיגדים "הייתי מעוניין ללמוד עם מרצה זה גם בקורסים אחרים" ו"הייתי ממליץ לאחרים ללמוד את הקורס אצל מרצה זה", אף כי היא מובהקת, מוסיפה כ-1% לשונות המוסברת.

◀ ברגרסיה על "תרומת הקורס להכשרה להוראה" אחוז השונות המוסברת הוא 79%. בולטת בעיקר תרומת המדד המבטא "מאפייני ידע פרקטי" בהוראת הקורס ($\beta = .67$), ומשתנה זה מסביר 76.6% מהשונות. שני המדדים האחרים במשוואה הם: המשתנה "הקורס תרם להרחבת הידע שלי בתחום שבו הוא עוסק" ($\beta = .223$), המוסיף 2.4% לשונות, והמדד "מאפייני ידע עיוניים-אקדמיים ששימשו בקורס" ($\beta = .040$), שתרומתו לשונות המוסברת, אף כי היא מובהקת ($p < .03$), היא מזערית וזניחה.

שתי משוואות רגרסיה אלה מעידות על שוני בתפיסת התרומה של לימודי ההכשרה להתפתחות האישית מחד גיסא ולהכשרה להוראה מאידך גיסא. אלה הן שתי תרומות שונות שהסטודנטים עשויים להפיק מלימודיהם.

סיכום הממצאים

- ◀ תפיסת הסטודנט את אופיו האפיסטמי של הקורס משפיעה על הערכתו אותו. מדד הקונצנזוס שחושב בכל קורס מראה, כי קיימת שונות רבה למדי בדיסציפלינות השונות הנלמדות במכללה (למעט בתחום החינוך המיוחד) בתפיסת הסטודנטים הלומדים באותו קורס את אופיו האפיסטמי. סטודנטים התופסים את הקורס כנושא אופי עיוני המנותק ממציאות ההוראה יטו להעריך אותו בציונים נמוכים מעמיתיהם התופסים את הקורס כמקושר למציאות ההוראה בין אם הוא נושא אופי עיוני ובין אם הוא נושא אופי מתודולוגי-דידקטי.
- ◀ נראה כי יש הרבה מהמשותף בין שני סוגי הקורסים שהסטודנטים מקשרים למציאות ההוראה. בשניהם הסטודנטים מתרשמים כי יש יותר ביטוי לשני בסיסי הידע, העיוני-אקדמי והפרקטי, וכן לשילוב ביניהם.
- ◀ ממצא פרדוקסלי המעיד על "אפקט המחיקה" התקבל בקורסים שהמעריכים תפסו את מאפייניהם האפיסטמיים כקורסים מטיפוס **עמנ**. סיווג הקורס ככזה התבטא גם בציוני הערכה נמוכים במדד הבוחן את מאפייני ה"שימוש בידע עיוני-אקדמי" בקורס.
- ◀ מניתוחי הרגרסיות על שני המשתנים התלויים, "תרומת הקורס להתפתחותי האישית" ו"תרומת הקורס להכשרתי להוראה", עולה ממצא מעניין שלפיו המדד המבטא את "איך" מתבצעת ההוראה לא נכנס למשוואות הרגרסיות.
- ◀ קיים "אפקט הילה" בהערכת ההוראה המתבטא בממצאים שלפיהם גם בהיבטים כלליים של הקורס, כמו: תרומתו להתפתחות האישית, תרומתו להרחבת ידע, ואפילו בהמלצה לאחרים להשתתף בקורס - סטודנטים התופסים את המהות האפיסטמית של הקורס כקורס מטיפוס **עמק** או **מתד**, נוטים להשתמש בציוני הערכה גבוהים במובהק מסטודנטים התופסים את הקורס כשייך לטיפוס **עמנ**. בהקשר של הערכת ההוראה, "אפקט ההילה" הוא ביטוי נוסף ל"אפקט המחיקה".

דין

הממצאים עונים בחיוב על שלוש שאלות המחקר שליוו את פיתוח השאלון להערכת ההוראה במטרה להאיר את מאפייניו במונחים של מהימנות ותוקף: (א) אכן, סטודנטים הלומדים באותו קורס ואצל אותו מרצה עשויים לתפוס את מאפייניו האפיסטמיים באופן שונה. (ב) הבדלים בתפיסה זו נוטים להתבטא גם בהבדלים בציוני הערכה

שהם נותנים להוראה ולקורס. מעריך התופס את הקורס כמנותק ממצייאות ההוראה, נוטה להעריך את ההוראה בקורס ואת הקורס עצמו בציונים נמוכים בהשוואה למעריך המתרשם שהקורס מקושר למצייאות ההוראה. באופן כזה עלולה להיווצר הטיה שיטתית בהערכת ההוראה והקורס. (ג) ההבחנה בין ממד ה"איך" לממד ה"מה" וכן הערכות כלליות נוספות שנבדקו מספקות מידע רלוונטי ומובחן על ההוראה, על הקורס ועל תפיסת הסטודנטים את תרומתו להכשרתם להוראה ולהתפתחותם האישית. נמצא שלא רק שהסטודנטים מבחינים בהערכותיהם בין שני ממדים אלה, אלא שממד ה"איך", המשקף היבטים דידיקטיים בתכנון הקורס ובהוראתו וכן את ההיבט של יחסי אנוש בין מרצה לתלמידיו, היבטים הנפוצים כל כך בשאלונים מסוג זה, כלל לא משמש גורם התורם לתפיסת הסטודנטים את תרומת הקורס להכשרתם להוראה וכן לא לתפיסתם את תרומתו להתפתחותם האישית. כלומר במונחים של בקרת איכות על ההוראה ועל מאפייני הקורס - מאפייניו האפיסטמיים של הקורס הם הרלוונטיים יותר בעיני הסטודנטים לשאלת תרומותיו להם.

המחקר מראה כי שימוש בשאלון להערכת ההוראה, שהנתונים שהוא מספק נחקרו ומגבלותיהם ידועות, מציג למעשה כלי הערכה מטיפוס הערכה מעצבת, המאפשר ללקוחות ההערכה להתייחס לנתונים תוך כדי בקרה על משמעותם. שימוש בממצאים כהערכה מסכמת מחייב להתייחס במלוא האחריות להטיה שיטתית הטבועה בנתונים, שמקורה בתפיסת המעריך את אופיו האפיסטמי של הקורס המוערך. לכן על הגוף המעריך חלה חובה לתקשר כהלכה את ממצאי ההערכה גם לגורמים המשתמשים בהם וגם למוערכים עצמם. יש להסביר למשתמשים בנתונים אלה את הגורמים המשפיעים עליהם ולספק להם, בנוסף לציוני ההערכה, את "מדד הקונצנזוס" לגבי אופיו הסכולסטי הנתפס של הקורס המוערך.

בסקירת הספרות הוצגה רשימה של הטיות המתקבלות משימוש בשאלוני הערכה בשיטה כמותית סגורה. אף אחד מהמחקרים לא בדק את הקשר בין תפיסת הסטודנט המעריך את מאפייניו האפיסטמיים של הקורס לבין הערכתו אותו. מחקרים שעסקו בקשרים כאלה השתמשו בהגדרות פורמליות לסוגי קורסים (קורסי מבוא, קורסי חובה, קורסי בחירה). העובדה שבאותו קורס לומדים סטודנטים התופסים את מאפייניו האפיסטמיים באופן שונה, ומשמעות עניין זה באשר להערכת ההוראה והקורס, הן נתון שלא נחקר עד כה באופן הזה. נתון זה משקף בעקיפין את מידת היכולת או הנכונות של סטודנטים להתמודד עם תוכני הקורס, את מידת עניינם בו ואת הקשרים השונים בין משתנים מתערבים אלה לבין תפיסתם את מאפייני ההוראה של המרצה. במילים אחרות, הערכת הסטודנטים את אופיו האפיסטמי

של הקורס נוטה "לצבוע" את הערכותיהם הנוגעות לאופן ניהול הקורס ואפילו את תפיסתם את מאפייני הידע המשמשים בו.

ממצאי המחקר עולים בקנה אחד עם ממצאי מחקרים המראים שסטודנטים נוטים לתת הערכות גבוהות יותר לקורסים שאופיים פרופסיונלי (Scriven, 1994) ולקורסים שהם מרגישים שלמדו בהם הכי הרבה (Wachtel, 1998). דפוס הערכה זה בולט יותר במכללות לחינוך בהשוואה לקורסים אוניברסיטאיים בדיסציפלינות מקבילות, שהלומדים בהם אינם מתעתדים לעסוק בהוראה (אבדור, 2000). התכלית העיקרית שהכשרת מורים מכוונת אליה, לפחות מנקודת ראותם של מרבית הסטודנטים, היא פיתוח מיומנויות בתחום ההוראה והחינוך. ואכן, הממצאים מראים כי בשאלת "תרומת הקורס להתפתחות האישית" - הגורמים העיקריים המסבירים תרומה זו קשורים להיבטים של הרחבת אופקים ושימוש בידע שמאפיינו עיוניים-אקדמיים. לעומת זאת בשאלת "תרומת הקורס להכשרה להוראה", שהיא תכלית הלימודים מבחינת הסטודנטים, הגורם העיקרי המסביר תרומה זו הוא היות הקורס בעל מאפיינים של ידע פרקטי (תכליתי), המכוון למציאות ההוראה ובית הספר ותורם להצלחה בתפקוד המקצועי.

תהליכי הוראה-למידה הם תהליכים מורכבים שאופיים הקשרי, וכל אחד מהמושגים המשמשים בתיאורם הוא מושג מורכב. מחקרים העוסקים בחקר תהליכי הוראה-למידה במוסדות להשכלה גבוהה מציגים מודלים שונים של יחסי גומלין ואינטראקציות בין משתנים הנגזרים משני משתני-על אלה, וביניהם משתנים הקשורים למאפיינים של מורים, סטודנטים, מבנה הדעת ומצבי הוראה-למידה. אחד המודלים המבוסס על ניתוח-על של ספרות מחקרית רחבה מראה כיצד ברמה של המרצה, כיחיד, האוריינטציה האישית שלו כלפי מהות הלמידה משפיעה על שיטות ההוראה שהוא נוקט, על מטלות הלמידה ומידת העומס שהוא יוצר בקורס ועל השיטה שבה הוא מעריך את הסטודנטים. משתנים הקשריים אלה חוזרים ומשפיעים על הגישות שהסטודנטים מפתחים כלפי הלמידה (Kember, 1997).

למעשה, ממצאי המחקר משקפים לא רק את מאפייני ההוראה והקורס, כפי שהסטודנטים תופסים אותם, אלא בעקיפין גם את רמות היכולת או הנכונות של סטודנטים בהכשרת מורים להתמודד עם תוכני הלימוד. המחקר מצטרף למחקרים העוסקים בהטיה המתקבלת בהערכות ההוראה, שמקורה בכמה גורמים החוברים יחד לכלל השפעה על מאפייני ההערכות הניתנות למרצה בקורס מסוים או באותו קורס שהוא מלמד בקבוצות שונות (Husbands, 1997). ממצא דומה ל"אפקט המחיקה" שנמצא במחקר הנוכחי נמצא גם במחקר אחר שנערך בקרב סטודנטים

הלומדים בתכניות להכשרת מורים במכללות לחינוך ובאוניברסיטאות. נמצא כי כאשר התבקשו הסטודנטים לציין את מידת חשיפתם לתכנים המבוססים על ידע עיוני-אקדמי, הם ציינו מידה נמוכה מאוד של שימוש בידע מסוג זה באותם קורסים עיוניים שנתפסו בעיניהם כמנותקים ממצייאות ההוראה (אבדור, 2001). "אפקט המחיקה" או "אפקט ההילה השלילית" מתגלה בהערכת קורסים שהסטודנטים אינם מוצאים עניין בתוכניהם העיוניים-אקדמיים. זהו מקור הטיה עיקרי שהמחקר הנוכחי מלמד עליו, הטיה שיש בה כדי לפגוע בתוקף הערכות הסטודנטים את ההוראה ואת הקורס. כלי המדידה, אף שהוא מהימן, אינו תקף (נבו, 2001). ההטיה אינה משפיעה על מהימנות הכלי, שכן היא שבה ומופיעה במדידות חוזרות באותם תנאים, אך היא משפיעה על תוקפו (על המידה שבה הכלי בודק את מה שהוא מתכוון לבדוק) ועל מידת ההוגנות בשימוש בממצאיו.

ממצאי מחקרים אלה מפנים את הזרקור לשאלות הנוגעות לקשר שבין מטרות המוצהרות בתכניות ההכשרה (הקוריקולום הגלוי), התרגום שלהן לתכנים ולדרכי הוראה-למידה המתרחשים בפועל (הקוריקולום הסמוי) והאופן שבו שני סוגי הקוריקולום משפיעים על מכלול אינטראקציות בין סגנונות הוראה של מורי מורים וציפיות שמפתחים סטודנטים כלפי הנלמד. חלק מהשפעת מכלול קשרים זה מתבטא בממצאי הערכת ההוראה. לא מן הנמנע שמרצים המבקשים לזכות בהערכות גבוהות מצד סטודנטים יוותרו על היבטים עיוניים-אקדמיים בהוראה לטובת גישות פרקטיות במובהק. שאלון הערכת ההוראה שפותח במחקר זה יכול לבחון כל קורס על פי רכיבים אלה ולהוות בסיס לבדיקת אופייה של ההוראה בקורס מול מטרותיו המוצהרות, ואף מול מטרות הקוריקוליום. תקצר היריעה מלעסוק בסוגיה זו, וגם אין זו מטרת המאמר. בספרות העוסקת בדגמים בהכשרת מורים (למשל: Britzman, Searle, & Pitt, 1997; Cochran-Smith, 1999; Elliott, 1993; Hativa, 1998; Kember & Gow, 1994; Murray & MacDonald, 1997; Pratt, 1997; Saroyan & Snell, 1997; Shulman, 1986). יש עיסוק רב בסוגיות אלה. מגמה בולטת בתחום הכשרת המורים, כמו במקצועות אחרים הנשענים על יסודות עיוניים-אקדמיים ופרקטיים כאחד, קוראת לשילוב מושכל בין שני ההיבטים.

הערכת ההוראה - בשביל מי ובשביל מה?

להליך הערכת ההוראה במוסדות להשכלה גבוהה יש כמה מטרות של בעלי עניין שונים שהם לקוחות ההערכה. לקוחות אלה צריכים להיות מודעים לבעיית ההטיה בתוצאות ההערכה, כפי שנמצאה במחקר זה כמו גם במחקרים אחרים.

מנקודת מבטה של הנהלת מוסד אקדמי

הערכת ההוראה באמצעות שאלוני משוב מסטודנטים היא תהליך שהנהלה האקדמית והאדמיניסטרטיבית יוזמת בדרך כלל מסיבות הקשורות לעבודה של ועדות מינויים וכוח אדם וכן כנתונים המשמשים את העוסקים בפיתוח סגל בתוך המוסד ובשיפור איכות ההוראה (Neumann, 2000). רשימה מעניינת של מחקרים שמצטט ווצ'טל (Wachtel, 1998) מראה, כי שימוש בתוצאות הערכת ההוראה אכן עוזר לשפר את ההוראה, וכי סטודנטים מעריכים בציונים גבוהים יותר את מי שלמדו ממנו יותר. אולם בה בעת המציאות מעידה על שימוש לא נכון בנתונים עד כדי פגיעה במוערכים בשל העדר מודעות לממצאי מחקרים המראים הטיות שיטתיות בהערכות אלה. בין ההבנות הנדרשות לצורך התייחסות אדמיניסטרטיבית ואקדמית להערכות ההוראה קיימת ההבנה, כי קיים הבדל בין הערכות ההוראה של סוגי קורסים שונים בתוך דיסציפלינות ובין דיסציפלינות, הנובע ממאפיינים אפיסטמיים של התכנים שבהם עוסקים. חשוב לא פחות לחקור מהם הקשרים בין היבטים שונים הנמדדים בשאלונים אלה לבין שאלות הקשורות לציפיות סטודנטים מהקורס, ליכולותיהם ולמידת עניינם להתמודד עם תכניו. חשוב גם להבין כי ניסוח ההיגדים בשאלונים מטיפוס זה, צורת סיכום תוצאות ההערכה ואופן העברת המידע משפיעים על תוכן הפרשנויות הניתנות לנתונים ועל משמעויותיהם ברמה אישית ומוסדית (Neumann, 2000).

מנקודת מבטו של סגל

אין בהכרח אהדה להליכי הערכת ההוראה מבחינת הסגל. מחקרים מראים כי בפקולטות שחויבו לקיים תהליכים של הערכת הוראה הייתה ירידה במורל הסגל, במידת שביעות רצונו מההוראה ובהיקף הדרישות בקורס ורמתן, ובכללן רמת המבחן. חוקרים טוענים כי מצב עניינים זה יכול להביא לאינפלציה בציונים. ממצאים אלה תומכים בהשערת "היד הרכה", שלפיה התקרבות לציפיות הסטודנטים ולדרישותיהם מחבבת עליהם את המרצים. בתנאים כאלה - הערכת ההוראה עלולה להיות אבן נגף במערכת ההשכלה הגבוהה, אם מרצים נכנעים ללחצים סמויים המופעלים עליהם

באמצעות שיטת הערכה זו, ומורידים מסיבות אישיות את רמת הדרישות ואת רמת ההוראה בקורס (Crumbly, Henry, & Kratchman, 2001; Wachtel, 1998).

מנקודת מבטם של סטודנטים

סטודנטים נענים ביתר אהדה למילוי שאלוני משוב על ההוראה אם הם עדים לתוצאות המבטאות שימוש בנתונים. מנקודת מבטם, הערכת ההוראה חשובה בראש ובראשונה לצורך שיפור ההוראה ובמקום השני לצורך שיפור בתוכני הקורס ובמבנה שלו (Yining & Hoshower, 2003). סטודנטים מצפים לפרסום תוצאות ההערכה כדי שיוכלו להשתמש בהן בשיקוליהם לבחירת קורסים (Babad, Darley, & Kaplowitz, 1999; Neumann, 2000). הם מעדיפים להעריך את הקורס באמצע סמסטר ולקבל תגובה מהמרצים על הערכה זו. הם גם מעדיפים שיקימו אתם ראיונות על פני שימוש בשאלוני משוב. בהערכת ההוראה מוצאים סטודנטים דרך "לשחרר קיטור", ולכן הם מאוכזבים כאשר אין תגובה אדמיניסטרטיבית להערכותיהם או כשאין שינוי בהתנהגות המרצים (Wachtel, 1998). באשר למידת ההוגנות של הסטודנטים בעת מילוי השאלונים, מחקרים מראים כי בצד דיווחים על הקדשת זמן ומחשבה הוגנים להערכה זו, יש גם סטודנטים המעידים על ניסיון להתנקם במרצה באמצעות ההליך (Crumbly, Henry, & Kratchman, 2001).

סיכום

כאמור, אחת הבעיות העיקריות במחקרי הערכת ההוראה היא העובדה שהערכה זו בדרך כלל נכפית על המרצים, אין בהכרח משקפת את תפיסת המרצה כאינדיווידואל או את תפיסתם של כלל חברי פקולטה באשר לטיבם של דרכי הוראה ושל חומרי למידה שרצוי ליישם כנגזר מתפיסתם את תכני הלימודים ואת מטרות הקורס. לשם כך דרוש **דיאלוג** בין מרצים לבין גורמים ממונים מבחינה אקדמית, בין מרצים בפקולטה לבין עצמם ובין מרצים לבין סטודנטים. דיאלוג כזה ישקף את אופייה של ההערכה כהערכה מעצבת, ועשוי לקדם מטרות המכוונות לאיכות אקדמית ופרופסיונלית של מרצים וסטודנטים כאחד. העדר דיאלוג מעמיד את המורה המוערך בעמדת נחיתות, כמי שמקבל ממצאי הערכה מסכמת העלולה לעתים גם לפגוע במעמדו האקדמי. לכך יכולה להיות השפעה בניסיון לשקם את הדימוי שלו תוך ויתור על סטנדרטים שהעמיד לעצמו כפרופסיונל וקיבוע דרכי הוראה המועדפות על המעריכים. בתנאים שבהם מתקיים דיאלוג פנים-מערכתי, ניתן להגיע תוך

משא ומתן להגדרות מקובלות יותר על הסגל באשר לאופייה של ההוראה ולדרכים לשיפורה. הדיאלוג מעצים את המשתתפים ומאפשר להם להתבונן בביקורתיות על הנחות "מובנות מאליהן" המכתיבות את המציאות היום-יומית של ההוראה (Gitlin & Smyth, 1990).

סוגיה זו מתקשרת להיבט חשוב מאוד בשאלת תוקף הממצאים המתקבלים בתהליכי הערכת ההוראה. גישות מודרניות לשאלת התוקף של שימוש בממצאי מבחנים חינוכיים או פסיכולוגיים מדגישות את אופן השימוש בממצאים. תפיסתו של מסיק (Messick, 1996, In: Brualdi, 1999; Orton, 1998) באשר ל"היבטים תוצאתיים של תוקף" (consequential aspects of validity) קונה אחיזה בקרב חוקרים רבים. במונח זה מבחין מסיק בין "תוקף תוצאה" לבין "תוקף מוכח בעליל", וטוען כי הבחנה זו נמצאת בבסיס שתי דרכים שבהן בחינה של בעיות פרקטיות עלולה אף היא להיות בעייתית. האחת - העדר ידע מספיק או הוכחות מספקות על מה שעולה ב"הוכחה מדעית" (כלומר בתוצאה, בממצאים), והאחרת, ולפיה מה שנחשב ל"הוכחה מדעית" ניתן לפרשנות בצורות שונות המשקפות נקודות מבט מוסריות שונות. הבחנה זו עוזרת למשתמשים בממצאים של מבחני הערכה למיניהם להיזהר מלהתייחס לבני אדם כאל חפצים בשעה שעל סמך ממצאי המבחן מתקבלות החלטות הרות חשיבות הנוגעות להם. "תוקף התוצאה" מצריך הוכחות והנמקות מסוג של "תוקף מוכח בעליל", כדי שניתן יהיה להעריך תוצאות צפויות ובלתי צפויות כאחת, העלולות לנבוע משימוש או מהעדר שימוש בממצאי המבחן. הערכת תוצאות אפשריות משני סוגים אלה חשובה במיוחד כאשר יש משמעות רבה לשימוש בממצאים של מבחני הערכה מבחינת היחיד או הקבוצה.

למרות הממצאים המעידים על בעיות של מהימנות ותוקף בהערכת ההוראה באמצעות שאלוני הערכה סגורים, אין בהם המלצה גורפת שלא לאסוף מידע כזה, בעיקר מפני שהתהליך זול ומהיר בהשוואה לתהליכים אחרים. גם לשיטות הערכה אלטרנטיביות שאופיין איכותני יש בעיות של תוקף ומהימנות, שאינן נופלות בהיקפן ובחומרן מאלו הקיימות בשיטות הכמותיות. הספרות העוסקת בהערכת תכניות מדגישה את הצורך בשימוש בכלי הערכה שונים הבודקים אותן מזוויות שונות, וכן את הצורך בבחינת הנתונים והממצאים המתקבלים מהערכות אלה בהקשרים שונים (Koon & Murray, 1995; Stufflebeam & Webster, 1980; Wright, 1990). באופן דומה יש להתייחס לנתונים המתקבלים משאלוני המשוב להערכת ההוראה כאל דרך אחת בבדיקת איכויות ההוראה והקורס, ויש להקפיד ולצרף אליה דרכי הערכה נוספות שישמשו בדיאלוג פנים-מוסדי בסוגיות אלה.

ביבליוגרפיה

אבדור, ש' (2000). ביטויים לתרבויות אקדמיות שונות המשתקפים בממצאי הערכת ההוראה וציפיות מהקורס בקרב סטודנטים במכללה להכשרת מורים וסטודנטים בקורסים אוניברסיטאיים. בתוך: המחקר בחינוך ויישומו בעולם משתנה, האגודה הישראלית לחקר החינוך (איל"ה) (עמ' 624-628). הכנס המדעי השנים-עשר. אבן יהודה: רכס.

אבדור, ש' (2001). אוניברסיטאות ומכללות לחינוך כסביבות להכשרה להוראה: מחקר משווה של בסיסי הידע וגישות חינוכיות המשתקפים בתפיסת פרחי הוראה את הכשרתם לבתי ספר על-יסודיים. חיבור לשם קבלת תואר "דוקטור לפילוסופיה", האוניברסיטה העברית בירושלים.

אבדור, ש' (2005). הערכת ההוראה במכללות לחינוך בממדי ה"איך" וה"מה" והשפעת תפיסת המעריכים את מאפייניו הסכולסטיים של הקורס על הערכתו. בתוך: ר' לידור, ב' פרסקו, מ' בן-פרץ ומ' זילברשטיין (עורכים), צמתים במחקר חינוכי. שיקולי דעת של חוקרים (עמ' 119-149). תל-אביב: מכון מופ"ת.

הר-זהב, ר' ומדינה, ב' (1999). דיני השכלה גבוהה. תל-אביב: הוצאת המחבר.

נבו, ד' (2001). הערכה בית ספרית: דיאלוג לשיפור בית הספר. אבן יהודה: רכס.

שילוח, י' ופרסקו, ב' (1994). תיקוף הערכה עצמית של מורה במכללה להכשרת מורים באמצעות הערכות תלמידיו - דו"ח מחקר. המכללה האקדמית בית ברל.

Airasian, P. W. (1997). *The classroom as an assessment environment* (ch. 1). MacGraw-Hill.

Aspinwall, K., Simkins, T., Wilkinson, J. F., McAuley, M. J. (1992). *Managing evaluation in education, a developmental approach* (pp. 3-5). London and New York: Routledge.

Babad, E., Darley, J. M., & Kaplowitz, H. (1999). Developmental aspects in students' course selection. *Journal of Educational Psychology*, 91(1), 157-168.

Barrett, J. (1986). *The evaluation of teachers*. (ERIC Document Reproduction Service, No. ED278657).

Bayer, A. E., & Braxton, J. M. (1998). The normative structure of community college teaching. *The Journal of Higher Education*, 69(2), 187-205.

Britzman, D. D., Searle, D., & Pitt, A. (1997). Toward an academic framework for thinking about teacher education. *Teaching Education*, 9(1), 15-26.

- Brualdi, A. (1999). *Traditional and modern concepts of validity*. ERIC/AE Digest. ED435714. (ERIC Clearinghouse on Assessment and Evaluation, Washington, DC).
- Burns, W. C. (1998). Colonizing the academy: Assessment, accountability and quality. *Educational Policy*, 12(4), 419-431.
- Cochran-Smith, M. (1999). *The future of teacher education: Framing the questions that matter*. Israel: Mofet Institute.
- Crumbly, L., Henry, B. K., & Kratchman, S. H. (2001). Students' perceptions of the evaluation of college teaching. *Quality Assurance in Education*, 9(4), 197-207.
- Darling-Hammond, L., Wise, A. E., & Pease, S. R. (1983). Teacher evaluation in the organizational context: A review of the literature. *Review of Educational Research*, 53(3), 285-328.
- Dunkin, M. J., & Barnes, J. (1986). Research on teaching in higher education. In: M. C. Wittrock (Ed.), *Handbook of research on teaching* (pp. 754-777). N.Y: Macmillan Publishing Company.
- Elliot, J. (1993). Three perspectives on coherence and continuity in teacher education. In: J. Elliot (Ed.), *Reconstructing teacher education*. London; New York: Falmer Press.
- Farrugia, C. (1996). A continuing professional development model for quality assurance in higher education. *Quality Assurance in Education*, 4(2), 28-43.
- Fresko, B., & Nasser, F. (2001). Interpreting student ratings: Conclusions, instructional modification, and attitudes towards course evaluation. *Studies in Educational Evaluation*, 27, 291-305.
- Gitlin, A., & Smyth, J. (1990). Toward educative form of teacher evaluation. *Educational Theory*, 40(1), 83-94.
- Hativa, N. (1998). Lack of clarity in university teaching: A case study. *Higher Education*, 36, 353-381.
- Husbands, C. T. (1997). Variations in students' evaluations of teachers lecturing in different courses on which they lecture: A study at the London school of economics and political sciences. *Higher Education*, 33(1), 51-69.
- Husbands, C. T., & Fosh, P. (1993). Students evaluation on teaching in higher education: Experiences from four European countries and some implications of the practice. *Assessment and Evaluation in Higher Education*, 18(2).

- Karapetrovic, S., Rajamani, D., & Willborn, W. W. (1999). University, Inc. *Quality Progress*, 32(5), 87-95.
- Kember, D. (1997). A reconceptualization of the research into university academics' conception of teaching. *Learning and Instruction*, 7(3), 255-275.
- Kember, D., & Gow, L. (1994). Orientations to teaching and their effect on the quality of student learning. *Journal of Higher Education*, 65(1).
- Koon, J., & Murray, H. G. (1995). Using multiple outcomes to validate students' ratings of overall teacher effectiveness. *Journal of Higher Education*, 66(1), Ohio State University Press.
- Long, P., Tricker, T., Rangelcroft, M., & Gilroy, P. (1999). Measuring the satisfaction gap: Education in the market-place. *Total Quality Management*, 10(4/5), 772-778.
- Murray, K., & MacDonald, R. (1997). The disjunction between lecturers' conceptions of teaching and their claimed educational practice. *Higher Education*, 33, 331-349.
- Neumann, R. (2000). Communicating student evaluation of teaching results: Rating interpretation guides (RIGs). *Assessment and Evaluation in Higher Education*, 25(2), 121-134.
- Nevo, D. (1994) How can teachers benefit from teacher evaluations? *Journal of Personnel Evaluation in Education*, 8, 109-117.
- Orton, R. E. (1998). Samuel Messick's consequential validity. *Philosophy of Education Yearbook*. PES Publications, U.S.A.
- Pratt, D. D. (1997). Reconceptualizing the evaluation of teaching in higher education. *Higher Education*, 34, 23-44.
- Ramsden, P., & Martin, E. (1996). Recognition of good university teaching: Policies from as Australian study. *Studies in Higher Education*, 21, 299-315.
- Robinson, R. K., Canty, A., & Fink, R. L. (1996). Public discourse of teaching evaluation: Privacy and liability considerations. *Journal of Education for Business*, 71(5), 284-289.
- Saroyan, A., & Snell, L. S. (1997). Variations in lecturing styles. *Higher Education*, 33, 85-104.
- Scriven, M. (1994). Using student rating in teacher education. *Evaluation Perspectives*, 4(1), 1-4.

- Scriven, M. (1995). *Student rating offer useful input to teacher evaluations*. (ERIC Document Reproduction Service, No. ED398240).
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4-14.
- Sirvanci, M. (1996). Are students the true customers of higher education? *Quality Progress*, 29(10), 99-102.
- Stark, J. S., & Lattuca, L. R. (1997). *Shaping the college curriculum*. Allyn & Bacon.
- Stringer, M., & Irwing, P. (1998). Students' evaluation of teaching effectiveness: A structural modeling approach. *British Journal of Educational Psychology*, 68, 409-426.
- Stufflebeam, D. L., & Webster, W. J. (1980). An analysis of alternative approaches to evaluation. In: J. S. Stark & A. Thomas (Eds.), *Assessment program evaluation* (1994, pp. 331-347). Boston: Pearson Custom.
- Wachtel, H. K. (1998). Student evaluation of college effectiveness: A brief review. *Assessment and Evaluation in Higher Education*, 23(2), 191-211.
- Wright, B. (1990). So how do we know it will work: An assessment memoir. In: J. S. Stark & A. Thomas (Eds.), *Assessment program evaluation* (1994, pp. 269-274). Boston: Pearson Custom.
- Yining, C., & Hoshower, L. B. (2003). Students' evaluation of teaching effectiveness: An assessment of student perception and motivation. *Assessment and Evaluation in Higher Education*, 28(1), 71-88.
- Zeichner, K. M., & Gore, J. M. (1990). Teacher socialization. In: W. R. Houston (Ed.), *Handbook of Research on Teacher Education* (pp. 329-348). New York: MacMillan.