

הערכה חלופית

שירה סופר-ויטל* ועידית פינקלשטיין**

מילות מפתח: הערכה חלופית, העידן הטכנולוגי, תוקף, מהימנות והוגנות בהערכה, ישרה אקדמית

<https://doi.org/10.54301/OHPA7380>

תפקידי הערכה מגוונים: (א) תפקיד מעצב של הלמידה - לצורך קידום ושיפור הלומד ומתן משוב; (ב) תפקיד מסכם - לצורך מדידה של הישגי הלומד וכאמצעי להכוונת הלומד למסלולי לימוד והכשרה התואמים את יכולתם, דיווח על טיב הביצוע וכמדד לקביעת קשיי למידה ואיתור הקושי בלמידה; (ג) תפקיד פסיכולוגי - להגברת מודעות ולהעלאת מוטיבציה וגם מתן חיזוקים חיוביים ושליילים ללומדים; (ד) תפקיד מנהלי - להפגנת סמכות ותיעוד הישגים (ניסן, 1985; Nevo, 1995).

בהערכה אנו מבחינים בין הערכה מסורתית לבין הערכה חלופית. המבחינים הם המוקד המרכזי של הערכה מסורתית, והמאפיין המרכזי שלהם הוא היותם אירועים בודדים ומבודדים, המשמשים כלי למדידת הישגי הלומד, לקבלת החלטות ביחס אליו ולמערכת בכללותה (למשל, מבחני מיצ"ב ומבחני פסיכומטרי) תוך התמקדות בציון הסופי (Nevo, 1995). נמצא כי השיטה המסורתית של מבחנים היא שיטת ההערכה השולטת כיום בבתי הספר ובהשכלה הגבוהה.

הערכה עשויה להתייחס לשני סוגים: הערכה מסכמת והערכה מעצבת. הערכה מסכמת כוללת מערכת של איסוף מידע או מבחן וצריכה לעמוד בשני קריטריונים: העברתה בסוף יחידת הוראה ושמירה על מדידה של הישגים לצורך קבלת החלטות (מתן ציונים, החלטה על קבלת תעודה, הסמכה, רישיון ועוד). הערכה מסכמת משקפת את הלמידה שכבר התרחשה ומעריכה אותה, והמוקד הוא שיפוט (לצורך מיון או הסמכה). הנחת היסוד של הערכה מסכמת היא ש"מידה אחת מתאימה לכולם", ולכן ההערכה תתבצע בהתאם לתוכנית אחת אחידה ומוקפדת (בירנבוים, 2018).

הערכה מעצבת - מורכבת ותהליכית. מטרתה של הערכה מעצבת היא לקדם את ההתפתחות של הלומד ולעודד למידה עצמית ואוטונומיה אקדמית (Brown & Knight, 1994) תוך התמקדות בהערכה דינמית (Black & Andrade & Cizek, 2010; Wiliam, 2009 S). מטרתה של ההערכה המעצבת היא עיצוב הלמידה והשפעה עליה תוך כדי התהליך. הערכה זו מספקת משוב לשיפור הלמידה וההוראה. רצוי שהמשוב יהיה ממוקד ומפורט, וזאת כדי לשקף את נקודות החוזקה והחולשה שלו. בהערכה זו יש מרכיב תקשורתי, המקדם למידה דיאלוגית. הערכה מעצבת עשויה לקדם תהליכי למידה במהלך הלמידה ותוך כדי הלמידה עצמה תוך שילוב בין הערכה ולמידה (הערכה לשם למידה, הל"ל). זאת בשונה מהערכה מסכמת שמטרתה היא מתן ציונים ללומדים (Birenbaum, 2016; Nevo, 1995).

בבתי הספר ובהשכלה הגבוהה השתרשה מזה זמן רב תרבות של הערכה ומתן ציונים, והם מרכיב חיוני, מכונן ויסודי של תהליכי הלמידה וההוראה בחברה המערבית. עם זאת, עולות תהיות ושאלות לגבי עצם ההערכה של ידע. למשל, מהו ידע? כיצד ידע יכול לבוא לידי ביטוי בתהליכי הוראה, למידה והערכה? מה ראוי וחשוב להעריך? כיצד ניתן להעריך? שאלות אלה מקבלות משנה תוקף בשל השינויים המהירים בטכנולוגיה, בחברה ובכלכלה, אשר התעצמו אף יותר בעקבות משבר הקורונה (Fitzmaurice & Ní Fhloinn, 2021). שינויים אלה מחייבים התאמה של מטרות החינוך במאה ה-21, ומימושן מחייב שינוי מהותי בהוראה ובהערכה.

הערכה חלופית מציעה לעבור ממוקד בהערכה כמדידה למיקוד באבחון הלמידה לשם שיפורה (בירנבוים, 2018). בהקשר של הערכה המתבססת על למידה שתהא משמעותית עבור הלומד, עולות שאלות חשובות, כגון האם וכיצד מעריכים נטיות ומיומנויות של יצירתיות, גמישות מחשבתית, חשיבה ביקורתית, תקשורת בין-אישית, חברתיות, ויסות רגשי, תחושת מסוגלות, שייכות ואוטונומיה? הערכה היא תיאור שיטתי של מושאים חינוכיים ושיפוט ערכם (Nevo, 1995). אך בשיפוט יש סובייקטיביות, ואילו התיאור יכול להיות אובייקטיבי למדי.

* ד"ר שירה סופר-ויטל, חברת סגל ומנהלת אקדמית בתואר השני בחוג לחינוך בקריה האקדמית אונו. לאחר סיום הדוקטורט במדעי הלמידה באוניברסיטה העברית, יצאה להתמחות פוסט-דוקטוריאלי בתוכנית The Israeli Centers of Research Excellence. מחקרה עוסקים ביצירת ידע שיתופי בקהילות עתירות טכנולוגיה, טיפוח למידה חברתית-רגשית ותהליכי למידה, הוראה והערכה בסביבות רב-תרבותיות.

** ד"ר עידית פינקלשטיין, מרצה בכירה ויו"ר ועדת ההוראה של הפקולטה למדעי הרוח והחברה בקריה האקדמית אונו. בעלת דוקטורט ביעוץ חינוכי ותואר ראשון במשפטים. תחומי המחקר שלה הם גולציה וחינוך, הוראה, למידה והערכה במוסדות חינוכיים, למידה רגשית חברתית רב-תרבותית.

מעידים בעיקר על ההבדלים ביכולות שנמדדו. בהערכה חלופית המהימנות תבוא לידי ביטוי גם ביכולות של הלומד ביחס לעצמו וגם ביחס לקבוצת השווים. תוקף (validity) בהערכה חלופית משמעו אם המשימות שהלומד מבצע משקפות את הידע ואת יכולותיו המשתנות כל העת. לעומת זאת, בהערכה מסכמת התוקף בודק את התאמת ציוני המבחן ליכולות או לידיע של הלומד ברגע נתון (פרידמן, 2012).

לפיכך הערכה חלופית היא תהליך שצוות חינוכי ולומדים משתמשים בו כדי לקבל משוב להתאמת הוראה ולמידה מתמשכת כדי לשפר את הישגי הלומדים בהתאמה למטרות של תוצאות הלמידה, ולכן היא מכונה גם הערכה מעצבת. לכן תוקף ומהימנות בהקשר של הערכה חלופית צריכים להיות מוטמעים בקוריקולום תוך אימוץ פדגוגיה ופרקטיקות הוראה מתאימות (McManus, 2008; Shepard, 2009; Otero-Saborido et al., 2018). אך נראה כי נדרש לפתח כלים ייעודיים לשם ההתייחסות לתוקף ולמהימנות בפרקטיקות של הערכה מעצבת (Otero-Saborido et al., 2018).

מדד חשוב נוסף בהערכה הוא ההוגנות (Fairness), המתייחסת למידה שבה המשימה ניתנת לכל הלומדים בתנאים אחידים ואינה יוצרת אפליה בין קבוצות מרקעים שונים. כלומר יש הקפדה על כך שלא תהיה הטיה בהערכת המשימה ובמתן הציונים. הוגנות המשימה משקפת את המידה שבה הלומדים תופסים את המשימה כהולמת את הסילבוס, הקוריקולום וההוראה בקורס. הוגנות קשורה לשוויון ולצדק בחינוך ובהערכה (Tierney, 2016). נראה כי ישנם כמה מרכיבים הדרושים לקידום הערכה חלופית הוגנת: (א) שוויון לימודי וצדק חברתי לכולם: הבנה של תחום התוכן תוך התאמה בין הקוריקולום לבין הערכה. נדרשת הבטחת איכותם, זמינותם ונגישותם של מכלול המשאבים הנחוצים ללמידה; (ב) מצע לימודי מכון, מאפשר ומיטיב, טיפוח אמון וכבוד בין הלומדים, עידוד הלומדים להשתתפות בלמידה; (ג) ביקורת ושקיפות של תהליך ההערכה: תכנון משימות תוך רפלקציה מתמדת ובחינת הנחות המשמשות להערכה. הערכה הוגנת צריכה להיות גלויה ושקופה.

בבתי הספר קיימת שתי תרבויות הערכה: תרבות הערכה ותרבות בחינה (בירנבוים, 2018), ואילו באקדמיה שולטת תרבות בחינה. לגבי הערכה בית ספרית קיים מחקר נרחב, אבל בהערכה באקדמיה ההיקף של המחקר מצומצם יותר (Finkelstein et al., 2022). ההערכה באקדמיה בחברה הישראלית צריכה לגלם מאפיינים תרבותיים שונים בין הסטודנטים המגיעים ממגזרים שונים. כלומר מגוון סטודנטים מגיעים עם תפיסות, אמונות, ציפיות, נטייה, ערכים, נורמות ורקעים חברתיים שונים. כאשר תרבות הבחינה היא השולטת, הציונים נתפסים באקדמיה כגורם מוכון פרופסיה, מסנן וממייין לפי יכולת סטטית של הסטודנטים. אולם סטודנטים שמגיעים ממגזרים שונים ומרקעים תרבותיים שונים בחברה הישראלית, מוערכים על פי אמות מידה

הערכה לשם למידה (ה"ל") (assessment for learning) היא גישת הערכה הקשרית ומעצבת באופייה. היא ממוקדת בלמידה ומטרתה היא חשיפת ההבנות של הלומדים ושיפורן תוך כדי הלמידה. כך תהליך הלמידה גלוי ושקוף, והצוות החינוכי בוחן ומעריך את הוראתו במטרה לקדמה (בירנבוים, 2018).

הערכה מעצבת היא חלופה להערכה המסורתית. היא

מתבססת על איסוף ראיות לצורך מתן פירוש משולב להישגי הלומד באמצעות משימות מורכבות ומאתגרות. בהערכה חלופית שמים דגש על ייחודיות הלומד, על השונות בקרב הלומדים וכן על קידום פיתוח מיומנויות מטה-קוגניטיביות (בירנבוים, 1997). יש חשיבות לשונות ולמגוון בין הלומדים. בתחומי הסייעוד, המדעים וההנדסה משתמשים בשיטת הערכה חלופית, כגון פרויקטים ומשימות אותנטיות (authentic assessment). הערכה מעצבת מייצגת מערכת מורכבת, המתאפיינת בהתמקדות בלמידה, והיא כוללת את הכיתה וקהילת הלמידה המקצועית אשר יוצרות כל העת ידע חדש. מלבד מיקוד בלמידה, מטרתה של הערכה זו היא לקדם את הנעת הלמידה, ליצור דיאלוג עם הלומד, להעצים את הלומד, לייחס ערך לשונות ולמגוון, לטפח חוללות עצמית וקולקטיבית (בירנבוים, 2018).

מהימנות, תוקף והוגנות בהערכה חלופית

מהימנות, תוקף והוגנות הם שלושת המדדים המקובלים לבדיקת איכות ההערכה (פרידמן, 2012). הספרות המחקרית מציעה כלים לבחינת מהימנות ותוקף של הערכה מסכמת (מבחנים). אנו רוצות להצביע על האתגר ועל הפער הקונספטואלי שבין מושגים הלוקחים מעולם התוכן של מדידה והערכה מסכמת לבין מושגים אלה בעת יישומם בעולם התוכן של הערכה חלופית. הנחת היסוד של הערכה חלופית היא ההתקדמות של הלומד, היכולת שלו להתפתח ולהשתנות. זאת בשונה מהערכה מסכמת אשר מודדת ומתייחסת למצב הסטטי של הלומד (Shepard, 2009). במאמר זה נצביע על התאמות אפשרויות לבחינת מושג המהימנות והתוקף באופן שיותאם להנחות היסוד של ההערכה החלופית.

מהימנות (reliability) בהערכה חלופית משמעותה בדיקת השיפור והלמידה של ציוני הלומד ביחס לעצמו לאורך זמן. המהימנות בוחנת אם הכלי אשר משמש להערכה משקף נאמנה את יכולותיו הדינמיות של הלומד. זאת בשונה מהערכה מסכמת שבה מהימנות משקפת את מידת העקביות והאפשרות לשחזר את תוצאות ההערכה. כלומר בהערכה מסכמת מהימנות מבקשת לבדוק אם תשובות הנבחנים לא ישתנו אם ינתן אותו כלי הערכה בזמן אחר (test-retest reliability), או באופן אחר (alternate form) או מעריך אחר (reliability between judges). מהימנות גבוהה משמעה שההבדלים בין ציוניהם של הנבחנים

הם המדד התקף, המהימן והאובייקטיבי למדידת הישגי הסטודנטים. הערכה חלופית שבנויה היטב יכולה לאפשר בקרה וניטור של רמת ההישגים. הערכה מסכמת אינה תמיד אינדיקציה מהימנה לאבחון רמת הידע של הסטודנטים, והתרומה הייחודית של הערכה חלופית בתהליך מעצב היא בכך שהיא מאפשרת לסגל ההוראה לצפות מראש את הפערים הלימודיים ולסייע למלא פערים אלה באמצעות שיטות הוראה והערכה חלופית (Tridane et al., 2015).

הערכה חלופית ככלי לקידום ישירה אקדמית ונשירה

הלמידה מרחוק מציבה אתגרים בפני אנשי חינוך. בלמידה פנים מול פנים אנשי חינוך יכולים לוודא השתתפות פיזית ונוכחות פעילה של הלומדים, ואילו בלמידה מקוונת הניטור ואפשרויות הבקרה על הנוכחות הצטמצמו מאוד. בלמידה מקוונת אנשי החינוך נעזרים בפתיחתן של מצלמות כאחד הכלים לבקרת נוכחות והשתתפות, וישנם מוסדות חינוך המחייבים פתיחת מצלמות. עם זאת, היכולת שלנו לבדוק השתתפות פעילה בשיעורים מצומצמת יותר מבעבר. טענתנו היא שככל שהניטור של המרצה הולך וקטן, אפשרות הנשירה של הלומדים הולכת וגדלה.

הערכה חלופית בתהליך מעצב עשויה לצמצם את תופעת הפלגיאריזם, ההעתקות, לתת מענה לסוגיית טוהר הבחינות ולקדם ישירה אקדמית. כיום ניתן לרכוש עבודות אקדמיות בנקל (גרוס-אביניר ומאירוביץ, 2021), ולכן עולה החשיבות של כלי הערכה חלופיים במוסדות חינוך. לכן הערכה חלופית בתהליך מעצב מקדמת את ההשתתפות הפעילה של הלומד לאורך כל תהליך הלמידה וכן עשוי להוביל למניעת העתקה ונשירה.

לסיכום, במאמר זה ניסינו לאתגר הנחות יסוד, כגון האם הערכה נתפסת בבתי ספר ובמוסדות אקדמיים כתהליך מבודד העומד בפני עצמו? האם נוכל להפיק מהערכה מעבר למדידת הרמה הנוכחית של הלומד, אלא גם לצורך קידומו? האם הערכה היא אבסולוטית? האם ניתן לאמוד את יכולותינו וידיעותינו על סמך ציון בודד של מבחן מסכם? דנו גם בחשיבות של היזון חוזר בין הערכה, פדגוגיה והוראה ולמידה. כמו כן, הצגנו את האפשרויות וההזדמנויות הקשורות להערכה חלופית, קידום השתתפות פעילה, התאמה לתמורות הטכנולוגיות, צמצום העתקה ונשירה.

המקורות

בירנבוים, מ' (1997). **חלופות בהערכת הישגים**. אוניברסיטת תל אביב, רמת.

אחידות, שלא תמיד תואמות את הרקע התרבותי והחברתי שלהם. אנו סבורות כי בחינה יכולה לשקף פחות התאמה למגוון הסטודנטיאלי מאשר ההערכה החלופית. הערכה חלופית גם בבתי ספר וגם באקדמיה עשויה להיות גשר בין התרבויות. הערכה זו גם מקדמת למידה, והיא גם תוכל לשמש גורם שמעריך ונותן ציון מהימן. הערכה חלופית בתהליך מעצב בוחנת את תהליך הלמידה לאורך זמן תוך כדי שהיא מקדמת את הלמידה של הסטודנט, מזהה את נקודות הקושי ומאבחנת אותן.

קידום הערכה חלופית בעידן הטכנולוגי

בעולם שבו מתרחשת מהפכה טכנולוגית נשאלות השאלות: האם בתי הספר והמוסדות להשכלה גבוהה מצליחים לעמוד בקצב השינויים המהירים? האם מוסדות אלה מצליחים להתאים את שיטות ההוראה, הלמידה וההערכה שלהם לשינויים הטכנולוגיים המהירים כפי שנדרש? בעולם שבו מתרחשת מהפכה טכנולוגית, השינויים הטכנולוגיים הם כה מהירים, ואנו חיים בעידן של הצפה במידע. הלומדים כיום ניזונים ממידע מהיר, זמין ונגיש, אך נדרשת מהם היכולת להבחין בין מידע מהימן ללא מהימן. מכאן, הולך וגובר הצורך לשנות את הדרך שבה אנו תופסים את תהליכי הלמידה, ההוראה ובייחוד ההערכה. הלומדים כיום ניזונים ממידע מהיר, זמין ונגיש. היכולת של כולנו לצרוך מידע הולכת ומשתכללת כל הזמן (McKnight et al., 2016) ולפיכך נדרשת התאמת ההערכה למאה ה-21.

מגפת הקורונה העצימה את הצורך בהתאמת תהליכי ההערכה באקדמיה לעידן של למידה והוראה מרחוק. לכאורה, הלמידה מרחוק וכניסתן המהירה של טכנולוגיות לאקדמיה עשויות להיות הזדמנות לקידום הערכה משמעותית (מג(ק)וונת (מקוונת ומגוונת). מחד גיסא, היכולת של כולנו לצרוך מידע הולכת וגדלה, ומאידך גיסא היכולת שלנו לוודא את מידת מהימנותו של המידע דורשת אימון, למידה וכן לעמוד על משמעות המידע, להפעיל לגביו חשיבה ביקורתית ואז לבצע סינתזה של המידע (Boursicot et al., 2020). למרות השינויים הטכנולוגיים המהירים והאפשרויות החדשות שנפתחות בפנינו כל העת בתחום ההערכה, נראה כי מוסדות חינוכיים עדיין דבקים בשיטות הערכה מסורתיות.

בבתי ספר אנו עדים להגוונה בשיטות ההערכה וכניסתן של שיטות הערכה חלופיות כשוות ערך לשיטות הערכה מסורתיות, כגון למידה מבוססת פרויקטים. כלומר בהתאם לגישות חינוכיות פרוגרסיביות, ישנן צורות ביטוי שונות של ידע. אפשר לייצג ידע בדרכים מגוונות באמצעות ביטויים מהאינטליגנציות השונות של גרדנר (Gardner & Hatch, 1989), דהיינו אפשר לבטא רעיונות בדרכי הבעה שונות.

אולם נראה כי התפיסה השולטת באקדמיה היא שהמבחנים

- Tierney, R. D. (2016). Fairness in educational assessment. In M. A. Peters (ed.), *Encyclopedia of Educational Philosophy and Theory*. Springer Science+Business Media. https://doi.org/10.1007/978-981-287-532-7_400-1
- Tridane, M., Belaouad, S., Benmokhtar, S., Gourja, B., & Radid, M. (2015). The impact of formative assessment on the learning process and the unreliability of the mark for the summative evaluation. *Procedia-Social and Behavioral Sciences*, 197, 680-685. <https://doi.org/10.1016/j.sbspro.2015.07.058>
- בירנבוים, מ' (2018). *חקר ויישום בהערכה לשם למידה (ה"ל) של תלמידים ומורים בישראל*. הוצאת מכון מופ"ת.
- גרוס-אביניר, א' ומאירוביץ, ע' (2021). התפתחות יושרה אקדמית של סטודנטים במהלך לימודי תואר. *סוגיות חברתיות בישראל*, גיליון 30 (2), 460-435.
- ניסן, מ' (1985). *צדק בכיתה - תפיסות מורים ותלמידים את העקרונות המדריכים מתן ציון*. בית-ספר לחינוך, 141-155.
- פרידמן, י' (2012). *סוגיות בהערכה תוקף במדידה ובמחקרי הערכה*. הוצאת מכון הנרייטה סולד.
- Andrade, H., & Cizek, G. J. (eds.). (2010). *Handbook of Formative Assessment*. Routledge. <https://doi.org/10.4324/9780203874851>
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability (formerly: Journal of Personnel Evaluation in Education)*, 21(1), 5-31. <https://doi.org/10.1007/s11092-008-9068-5>
- Boursicot, K., Kemp, S., Ong, T. H., Wijaya, L., Goh, S. H., Freeman, K., Curran, I. (2020). Conducting a high-stakes OSCE in a COVID-19 environment. *MedEdPublish* 9 (1). <https://doi.org/10.15694/mep.2020.000054.1>
- Finkelstein, I., Soffer-Vital, S., Shraga-Roitman, Y., Cohen-Liverant, R., & Grabelsky-Lichtman, Z. (2022). An integrative multi-dimensional model of culturally relevant academic evaluation for the 21st century. *International Journal of Higher Education*, 11(1), 187-200. <https://doi.org/10.5430/ijhe.v11n1p187>
- Fitzmaurice, O., & Ní Fhloinn, E. (2021). Alternative Mathematics Assessment During University Closures Due to Covid-19. *Irish Educational Studies*, 40 (2), 187-195. <https://doi.org/10.1080/03323315.2021.1916556>
- Gardner, H., & Hatch, T. (1989). Educational implications of the theory of multiple intelligences. *Educational Researcher* 18 (8), 4-10. <https://doi.org/10.3102/0013189X018008004>
- McManus, S. (Ed.) (2008). *Attributes of effective formative assessment*. Washington.
- McKnight, K., O'Malley, K., Ruzic, R., Horsley, M. K., Franey, J. J., & Bassett, K. (2016). Teaching in a digital age: How educators use technology to improve student learning. *Journal of research on technology in education*, 48(3), 194-211. <https://doi.org/10.1080/15391523.2016.1175856>
- Nevo, D. (1995). *School-Based Evaluation: A Dialogue for School Improvement*. Pergamon University Press.
- Otero-Saborido, F. M., Sánchez-Oliver, A. J., Grimaldi-Puyana, M., & Álvarez-García, J. (2018). *Flipped learning and formative evaluation in higher education*. Education+ Training. <https://doi.org/10.1108/ET-12-2017-0208>
- Shepard, L. A. (2009). Commentary: Evaluating the validity of formative and interim assessment. *Educational Measurement: Issues and Practice*, 28(3), 32-37. <https://doi.org/10.1111/j.1745-3992.2009.00152.x>