

## האמינות של ההערכה

### בשיטת החלופות בהערכת הישגים [1]

#### שני מדדי האמינות

כערך מילוני מוגדר המונח **אמינות** כמאפיין של אדם או של מקור מידע, שאפשר לבטוח בהם או לסמוך עליהם [2]. בזיקה לציונים המופקים מכלי הערכה שונים השאלה היא: באיזו מידה הציונים המוגדרים שבהם מדובר אכן משקפים נאמנה את המאפיינים המבוקשים של הלומדים המוערכים הנוגעים בדבר?

בנוגע לאמינות נשמעים לעתים היגדי ביקורת בבתי הספר. ביקורתם של תלמידים מופנית כלפי אמינותן של מטלות הערכה בית ספריות, וביקורתם של מורים מופנית, לעתים, כלפי אמינותן של בחינות חיצוניות. אלה דוגמאות של היגדי ביקורת:

1. **תלמידים**, בסיום שנת לימודים, בקשר למכלול כלי ההערכה:  
"למדנו הרבה נושאים ורכשנו גם מיומנויות – איך ללמוד טוב. אבל בבחינות היה קשר רק לחלק קטן ממה שלמדנו. יש הרבה נושאים חשובים שלמדנו ועליהם לא הייתה בכלל הערכה."  
2. **תלמידים** – בקשר לבחינה:  
" במבחן הזה רק חלק מהשאלות קשורות למה שלמדנו, רוב השאלות לא שייכות לחומר הזה. איך אנחנו יכולים לענות על שאלות בנושאים שלא למדנו?"
3. **תלמיד** – לרכז מקצוע – בקשר לפריט מסוים בבחינה:  
" המורה סימן תשובה שגויה והעיר: 'התשובה שלך איננה תשובה לשאלה הזאת'. אני לא מסכים, עניתי בדיוק על מה שהבנתי בשאלה. הנה, תקרא אתה את השאלה ותגיד לי: איך אתה מבין את השאלה? כנראה שבראש של המורה יש שאלה מסוימת, אז שישאל בדיוק את מה שהוא מתכוון".
4. **מורים לאנגלית** – בנוגע למטלת "קטע שלא נלמד" (unseen) – בבחינת בגרות 5 יח"ל:  
"השאלות בקשר לקטע אינן מתחום הבנת הנקרא (comprehension), אלה שאלות היגיון ולוגיקה".
5. **מורים בישיבת צוות** – שניים מן המורים בצוות הכינו הצעה לשאלון של בחינה שכבתית. בנוגע לכל אחד מן הפריטים נקבעו המאפיינים (בטבלת מתאר דו ממדית - ראו נספח לשאלון בחינה ב"קטע שלא נלמד" בגמרא): ממד אחד הוא הנושאים התוכניים וממד שני הוא המטרות/ המיומנות הנבדקות באמצעות הפריטים השונים.  
מורים מן הצוות העירו ביחס לשלושה מתוך עשרת הפריטים:

"הקשר בין ניסוח הפריט ובין ציון מאפייניו איננו חד-חד ערכי. נדרש שינוי הפריט כדי להתאימו למאפייניו או לחילופין לשנות את המאפיינים על מנת שיתאימו לנוסח הפריט".

6. **תלמיד** – בנוגע לציון על עבודת חקר שהכין:

ערערתני לפני המורה ורכז המקצוע על הציון שקיבלתי. אמרתני להם שבבדיקה עצמית שלי על פי המחונן שהמורה נתן לנו – ראיתי שצריך להוסיף לציון שלי 10 נקודות. אכן, אחרי בדיקה שנייה הוסיפו לי 6 נק'.  
רכז המקצוע הסביר לי את התהליך שהיה: העבודה נמסרה למורה אחר מצוות המקצוע, שהעריך את העבודה שלי על פי המחונן וקבע שצריך להוסיף לי 8 נק'. שתי ההערכות, הערכת המורה שלי וההערכה של המורה הנוסף הוצגו בפני צוות מורי המקצוע. בדיון שהיה בצוות סוכמה תוספת של 6 נקודות.  
ידוע לנו, התלמידים, שלצוות המורים יש ישיבות קבועות. בישיבות האלה הצוות מכין את הבחינות ואת מטלות ההערכה של השכבה. הצוות גם בודק במשותף את התהליך של קביעת הציונים.

7. **רכז מקצוע** – הביע תמיהה לאחר ששני מורים העריכו מבחן של תלמיד באופן בלתי תלוי, ופער ההערכות בין השניים הגיע ל – 20 נק'. הרכז הביא את הנושא לדיון בצוות מורי המקצוע.

8. **תלמיד** – נכשל בבחינת הבגרות במועד הקיץ. התלמיד החליט להמר, הוא חזר וניגש לבחינה במועד ב' ללא הכנה נוספת. במועד זה הוא עבר את הבחינה בהצלחה בציון שהיה גבוה ב 10 נק' מהציון במועד הקיץ. באוזני חבריו התפאר התלמיד בהישג שאותו השיג ללא מאמץ מיוחד. **תלמידים** הפנו שאלה למורה: "אם אכן התלמיד השיג שני ציונים שונים (כישלון מול הצלחה) – כשהוא מצויד באותה רמת ידע – מהו, אפוא, הציון האמיתי שלו?"

בשמונת ההיגדים האלה הביקורת היא כלפי שני מדדי האמינות של הערכת ההישגים: **התקפות** ו**המהימנות**. בהיגדים 1 - 5 הטיעונים הם מתחום תקפות התוכן של ההערכה דהיינו, באיזו מידה אכן מקיפה הערכת ההישגים מדגם מייצג של כלל התכנים, ושל מטרות ההוראה והמיומנויות שבנוגע אליהם נדרשת הערכה. בהיגדים 6 – 8 הטיעונים הם מתחום המהימנות של ההערכה, דהיינו, בדיוק ובמידת האובייקטיביות של הערכת ההישגים והציונים המופקים ממנה. לפירוט יתר ולצורך התאמה של שני מדדי ההערכה למתכונת הערכה המבוססת על שיטת החלופות בהערכת ההישגים מובאים להלן נוסח מוצע להגדרה של כל אחד משני המדדים האלה בצירוף דברי הסבר.

**תקפות התוכן** – מבין שלושה סוגי תקפות המאוזכרים בספרות המקצועית הרלוונטית: תוקף מושג נבנה, תוקף ניבוי ותוקף תוכן [3] ההתייחסות במאמר הזה היא לתוקף התוכן דהיינו למידה שעל פיה, באמצעות כל אחד מכלי ההערכה בפרט, ובאמצעות כלל כלי ההערכה אכן מושגת הערכת ההישגים של מה שמתכוונים להעריך (בנושא הזה ראו גם את [המצגת תקפות](#)).  
**מהימנות** - רמת הדיוק ומידת העקיבות של תהליך הערכת ההישגים. זאת, תוך התייחסות לכל אחד מכלי ההערכה ולשילובם יחדיו (בנושא הזה ראו גם את [המצגת מהימנות](#)).

בהמשך הדברים מפורטים דברי ההסבר בנוגע לכל אחד משני המדדים האלה: תקפות ומהימנות תוך התייחסות למקומן בשיטת החלופות בהערכת הישגים.

## **תקפות**

בנוסח ההגדרה שהוצע למושג "תקפות התוכן" מדובר על הערכה של "מה שמבקשים להעריך". השאלה הנשאלת היא: מהו אותו עולם תוכן שאותו מבקשים להעריך? על פי האמור להלן ההתייחסות היא לשני היבטים כדלקמן.

### 1. הערכת הישגים פרטנית

הערכת הישגים פרטנית נעשית על פי כל אחד מכלי ההערכה. בהקשר הזה עניינה של התקפות הוא במידת ההתבססות על הוכחות טובות ומספיקות. זאת, מבחינת פרשנות ההערכה ומבחינת המסקנות המופקות מן הציונים בהקשר לכלי ההערכה הספציפי.

### 2. הערכת הישגים כוללת

הערכת הישגים כוללת מבוססת על תכנית הלימודים של מקצוע מסוים ומידת ביצועה בפועל. בהקשר הזה נדרשת תשומת לב לזיקה מרבית בין תכנית הלימודים לבין הערכת הישגים במקצוע הנלמד. לפיכך, תקפות הערכת הישגים גדלה ככל שהיא מייצגת, הן את התכנים והן את הטווח הרחב של המטרות בתחום החשיבה, בתחום מיומנויות הלמידה העצמית והשיתופית, ואף בתחום החווייתי-רגשי. הכוללות בהערכת הישגים ניתנת להשגה באמצעות מכלול של כלי הערכה מתאימים.

## **מהימנות**

כאמור, במושג "מהימנות" ההתייחסות היא לרמת ההגינות האובייקטיבית כלומר לדיוק ולמידת העקיבות בתהליך הערכת הישגים, רמה זו חיונית בקביעת ציוני הלומדים. מצב של שונות בקביעת הציון בתהליך של הערכת הישגים חוזרת מוכר למורים ואף לתלמידים. המקורות לשונות בהערכת הישגים חוזרת ניתנים למיון לשני סוגים:

### 1. שונות התלויה במוערך

זהו מצב אפשרי שבו המוערך חוזר על ביצוע מטלה שוות ערך בהפרש זמן מבלי שחל שינוי ברמת הידע והשליטה שלו בתחום הנבדק, ולמרות זאת, איכות ביצועו ובעקבותיה קביעת ציונו משתנים. השונות בקביעת הציון עשויה להיות מוסברת בשינויים במצב הרגשי של המוערך או במצבו הגופני, וכן – בניסוחי מטלות שאינם בהירים די הצורך. ניסוח מטלה שאינו בהיר עשוי להיות מובן באופן שונה על ידי המוערך בכל אחת משתי הפעמים שבהן הוא מתמודד עם המטלה.

### 2. שונות התלויה במעריך

זהו אופן הערכה שבו שתי הערכות או יותר מתבצעות במקביל על ידי מעריכים שונים או על ידי אותו מעריך בהפרש זמן. כשההערכה היא של מעריכים שונים, עשויה השונות בהערכה חוזרת להיות מוסברת בהתבססות על תבחיני הערכה עמומים (או על אי קיום תבחינים כלל). במקרה של הערכה נוספת על ידי אותו מעריך, ההסבר נעוץ גם בשינויים במצבו הרגשי או הגופני של המעריך (למשל, שעמום או עייפות).

בהמשך הדברים ההתייחסות היא בעיקר לסוג השני של הערכה חוזרת: ביצוע הערכות אחדות (שתיים לפחות) דהיינו 'מהימנות בין מעריכים', זהו סוג הערכה רונח ואף מומלץ לביצוע. בהקשר

הזה ראוי לציין שכל הערכה המלווה בציון איננה אלא אומדן של רמת הידע האמיתי. השאיפה בתהליך ההערכה היא לצמצם, ככל האפשר, את הפער של טעות המדידה בין הידע האמיתי לבין הציון שנקבע, שהוא הציון הנצפה [4]. בחלופות בהערכת ההישגים ההנחה היא ששקלול רצף ההערכות, תוך שמירה על המשכיותן, עשוי אף הוא לתרום לצמצום הפער הזה.

### **מה בין תקפות למהימנות?**

בספרות המקצועית הנוגעת בדבר מובהר שהחלופות בהערכת ההישגים התפתחו על רקע הביקורת שהועלתה כלפי הערכת ההישגים המסורתית וכמובא להלן. אחד הגורמים המרכזיים שהביאו להתפתחות המואצת של החלופות בהערכת ההישגים הוא הטיעון, שרמת התקפות היחסית בהערכת ההישגים המסורתית היא נמוכה. ההערכה המסורתית מבוססת, בדרך כלל, על שימוש בכלי הערכה יחיד, מבחן. לפיכך, האפשרות להעריך את מלוא תכנית הלימודים וביצועה בפועל היא מוגבלת. כמו כן, גם בעת השימוש בכלי הערכה יחיד זה – במבחן – יש מקרים שבהם ניסוח מטלות הבחינה מצמצם את טווח הבדיקה: הניסוחים מכוונים לבדיקת מטרות בתחום הידע בלבד. בעיקר, במבחנים מסוג "רבי בחירה", שבהם הפריטים הם פריטים סגורים. זאת, למרות העובדה שניסוח מתאים של פריטים מסוג זה עשוי לאפשר בדיקה של תפקודי חשיבה גבוהים גם במבחנים מסוג "רבי בחירה".

לעתים נראה, שהגדלת התקפות מצריכה ויתור על רמת המהימנות. זאת משום שמהימנות גבוהה ואובייקטיביות מרבית ניתנות להשגה אם מטלות ההערכה סגורות וממוקדות. אין קושי להעריך מטלות מסוג זה תוך הקפדה על אובייקטיביות מרבית, משום כך המבחנים "רבי הבחירה" מכוונים גם "מבחנים אובייקטיביים". השימוש במטלות הערכה פתוחות המזמינות התייחסות רב כיוונית, ולא דווקא חד משמעית, מאפיין את המטלות המאפשרות להעריך כישורי חשיבה מסדר גבוה. כמו כן, השאיפה בחלופות בהערכה היא להגדיל ככל האפשר את טווח מטרות ההוראה הנתונות להערכה, בתחומי החשיבה, הרגש והתקשורת הבין-אישית וברמות ההישג השונות: מהרמה הנמוכה לרמה הגבוהה, מטלות מסוג זה קשה להעריך. לפיכך, ככל שגדלה תקפות ההערכה כך עלולה להצטמצם מהימנותה. לעניין זה נשאלת השאלה: כיצד ניתן להגדיל את רמת התקפות של הערכת ההישגים תוך פגיעה מזערית, ככל האפשר, ברמת המהימנות שלה? תשובה לשאלה זו ניתן למצוא בעבודה השיתופית של צוות מורי המקצוע.

### **הבטחת אמינות ההערכה כחלק מהעבודה השיתופית של הצוות**

כאמור, תנאי הכרחי (אך בלתי מספיק) להבטחת האמינות של החלופות בהערכת ההישגים היא העבודה השיתופית של צוות מורי המקצוע. ההתייחסות לסוגיות הקשורות לתקפות ולמהימנות של ההערכה היא מן הנושאים המרכזיים בישיבות הצוות. בהקשר הזה דיוני הצוות נושאים אופי של "שקלא וטריא" ומיועדים לבירור יסודי ומעמיק של הסוגיות האלה. להלן הצעה של תהליך לבירור מידת האמינות דהיינו, התקפות והמהימנות של כלי ההערכה במסגרת עבודת הצוות.

## התקפות

לשם חיזוק התקפות ביישום החלופות בהערכת ההישגים ראויים לתשומת לב שניים מן העקרונות המנחים את מהלכי הביצוע של הערכת ההישגים, ואלה הם :

**רב שיטתיות:** הערכת ההישגים אמורה להקיף את התפקוד של הלומדים בתחום ההכרה (החשיבה) ובתחום הרגש (עמדות וערכים), ובכל אחד משני התחומים האלה את כל רמות ההישג: מהרמה ההתחלתית עד לרמה הגבוהה. לצורך הזה ההערכה מתבצעת בשיטות מגוונות תוך הסתייעות בכלים שונים: בחינות בכתב, ביצועים בעל פה, תלקיטים, עבודות חקר, מטלות ביצוע, מפות מושגים, מטלות קבוצתיות ועוד.

**רצף והמשכיות:** הערכת ההישגים היא רצופה ומצטברת, בכך יש כאן הימנעות מהערכות שבעיקרן הן חד פעמיות ומשום כך נפגמת אמינות ההערכה. עם זאת, יש להימנע מריבוי יחסי של מצבי הערכה.

נושא התקפות נוגע לשני היבטים :

1. הערכת הישגים כוללת בזיקה לתכנית הלימודים.
  2. הערכת הישגים פרטנית בזיקה לכל אחד מכלי ההערכה. לפיכך במהלך דיוני הצוות יתבצע התיקוף תוך התייחסות לכל אחד משני ההיבטים האלה כמובא להלן:
- א. **התייחסות להערכת ההישגים הכוללת** - התייחסות כוללת לתקפות מצריכה פריסת תכנית לימודים של המקצוע, גזירת תכנית הוראה על פיה ומיפוי מגוון כלי הערכה מול סעיפי התכנית.
- השאלות שמן הראוי לברר הן אלה :
- מהן מטרות ההוראה, שאת מידת השגתן מבקש הצוות להעריך?
  - מהן מיומנויות הלמידה שאת מידת השליטה בהן מבקש הצוות להעריך?
  - אילו תכנים ייבדקו?
  - אילו כלי הערכה יפותחו לצורך בדיקת מירב המטרות, המיומנויות והתכנים? רצוי שבירור השאלה הזאת ייעשה בזיקה למיפוי שנתי ורב שנתי, ואף למיפוי בזיקה לכלל מקצועות הלימוד. זהו תהליך המאפשר שימוש מתואם במגוון כלי ההערכה ובמקצועות הלימוד השונים.
  - כיצד תובטח השמירה על טוהר המידות בתהליך ביצוע המטלות תוך התייחסות לבעיות ולקשיים שלהלן?
    - הבעייתיות של טוהר המידות עלולה להחריף בהקשר של החלופות בהערכה. זאת, בגלל הסיבות האלה:
      - חלק ניכר מן המטלות מתבצע מחוץ לכותלי הכיתה, ובמשכי זמן ארוכים באופן יחסי.
      - לעתים ניתן עידוד לתלמידים להסתייע בגורמי עזר, אך בתנאי שישמרו עצמאותם ומקוריותם של המוערכים.
      - מקורות המידע האלקטרוניים עלולים להפוך ממשאבי מידע למקור לעבודות מוכנות.
      - בפעילות קבוצתית קשה לעתים לזהות את התרומה של כל אחד ואחת מן התלמידים השותפים לביצוע המטלה.

**ב. התייחסות פרטנית לכל אחד מכלי ההערכה – התיקוף הפרטני הוא תהליך שעל-פיו**

קובעים שרמת התוקף של כלי ההערכה היא גבוהה. התיקוף של כל אחד מכלי ההערכה ייעשה על ידי צוות מורי המקצוע. זאת, בעקבות הגדרה של תחומי בדיקת הכלי. להגדרת התחומים יסתייע הצוות בטבלת מתאר דו-ממדית או בדרך מיפוי אחרת. צוות המורים יתייחס לארבע השאלות האלה:

- באיזו מידה אכן מייצג תוכן הכלי באמצעות מדגם מייצג של פריטים את מגוון הנושאים, המטרות והמיומנויות, שאת מידת השגתן נועד הכלי להעריך?
- האמנם הניסוחים הכלליים (ההנחיות לביצוע המטלה) ואף הניסוח של כל אחד מן הפריטים, שקופים וברורים במידה שהתלמידים מבצעי המטלה אכן מבינים את אשר נדרש מהם?
- בזיקה לכל אחת מן המטלות ולכל אחד מן הפריטים: האמנם הקשר בין הגדרת התחום הנבדק ובין ניסוחו בפועל הוא חד-משמעי?
- האם הפריטים האלה, ולא אחרים, הם הפריטים שעליהם ראוי לבסס את ההערכה?

זוהי טבלת סיכום של שני היבטי התייחסות לתקפות:

**שני היבטים של התקפות**

<b>הערכת הישגים פרטנית באמצעות כלי הערכה אחד</b>	<b>הערכת הישגים כוללת באמצעות מגוון כלי הערכה</b>
<ul style="list-style-type: none"><li>• כיסוי מתאים של הקטע הרלוונטי בתכנית הלימודים, שאת מידת השגתו מבקשים להעריך באמצעות הכלי.</li><li>• ניסוח ברור של ההנחיות בקשר לאופן השימוש בכלי, וניסוח בהיר של כל אחד מהפריטים ומן המטלות.</li><li>• בזיקה לכל אחד מן הפריטים ולכל אחת מן המטלות: הבטחת קשר חד משמעי בין הגדרת התחום הנבדק ובין ניסוח הפריט או המטלה.</li></ul>	<ul style="list-style-type: none"><li>• ייצוג מדגמי של מכלול תכנית הלימודים ומימושה בפועל.</li><li>• הצבת כלי ההערכה השונים מול מפרט הנושאים, המטרות והמיומנויות.</li></ul>

התיקוף איננו מסתיים בשלב הפיתוח של כלי ההערכה. יש המשך לתהליך הזה גם לאחר ביצוע המטלות. אלה האפשרויות בשלב הזה:

- ניתוח תגובות רפלקטיביות של התלמידים בקשר לתפיסתם את איכות המטלות.
- ניתוח פריטים שרמת הישגים היחסית בהם היא נמוכה. בזיקה לפריטים האלה תבצע בדיקה: באיזו מידה קשורה התוצאה הזאת בקשיים הכרוכים ברמת התקפות של ההערכה?

## המהימנות

חיזוק רמת המהימנות של ההערכה במסגרת עבודת הצוות ניתן להשגה בדרכים האלה :

- פיתוח מחוונים לכלי ההערכה השונים. זאת, תוך הקפדה על שקיפות מרבית שלהם בהינתנם בידי מעריכים שונים : מורה הכיתה, מורים עמיתים, משקיפים מבחוץ ואף תלמידים בתפקודם כמעריכים בתהליך של הערכה עצמית. (ראו דוגמא של [טופס ציינון](#) במסגרת דברי ההסבר על המבחן).
- אימון המורים בהערכה על פי המחווון : בחירה אקראית של תוצר תלמיד שאותו מעריך, באופן בלתי תלוי, כל אחד מחברי הצוות. מתקיים ברור של קשיים ומתקיים דיון במגמה להגיע להסכמות.
- התהליך של קביעת "מהימנות בין מעריכים" כרוך בבדיקת תוצרי התלמידים לצורך ציינון. בהקשר הזה מוצע התהליך כדלקמן :  
טרם הבדיקה של כלי הערכה מסוים (מבחן ; שיקוף פרק לימוד ; מטלת ביצוע) המורה דוגם שלוש עבודות של תלמידים (בקבוצת לימוד שמעל 15 תלמידים). הדגימה היא אקראית של אומדן הישג התחלתי, בינוני וגבוה. שלוש העבודות נבדקות באופן בלתי תלוי על-ידי המורה ועל-ידי מורה עמית. זאת, תוך שימוש בטופס הציינון. בהשוואת התוצאות של שתי הבדיקות מבררים פערי הערכה ומגיעים להסכמות (מקרים מיוחדים של חוסר הסכמה מובאים לדיון נוסף בצוות מורי המקצוע). בעקבות ביצוע הליך המהימנות בין מעריכים ולאחר אימוץ ההסכמות שהגיעו אליהן שני המעריכים בודק המורה את כלל העבודות ומציינן אותן.

## סיכום

רמת האמינות בהערכת הישגים מותנית בשני המדדים : מדד התקפות ומדד המהימנות. עניינה של התקפות הוא בייצוג מדגמי מרבי של התכנים ושל מטרות ההוראה, ובהבטחת מצב שבו ממצאי ההערכה אכן מבוססים על נתונים מוצקים. משמעה של המהימנות הוא הדיוק, ההוגנות והעקיבות בהערכה.  
החלופות בהערכת הישגים מזמנות אפשרויות לחיזוק שני מדדי האמינות האלה. תנאי הכרחי לכך הוא עבודה שיתופית שיטתית ומסודרת של צוות מורי המקצוע.

## מקורות

[1] עיבוד של מאמר שראה אור ב"מדריך מעשי לחלופות בהערכת הישגים" : "איך עושים את זה?" (2004) עורכים : רז רות ובן-אליהו שלמה, משרד החינוך, ירושלים, עמ' 48 – 58. ב"מדריך" מפורטות שיטות הוראה והערכת הישגים, וכן – מוצגים כלי הערכה שפותחו בפרויקט הניסויי "בגרות 2000" (1995 – 2000).

[2] ראו : מילון אבן שושן 1975 ; שויקה 1997

[3] פרידמן, יצחק (2005), **מדידה והערכה של תכניות חברתיות וחינוכיות**. מכון הנרייטה סאלד, ירושלים, עמ' 231 – 233

[4] פרידמן, כנ"ל, בנושא : "המהימנות כתוצר של התאוריה הקלסית של המדידה", עמ' 219 -