

פיתוח מטלת הכתיבה בתחום המילולי של מכפ"ל

גלעד סופר
ספי פומפיאן
נעמי גפני

מאי 2013



מרכז ארצי לבחינות ולהערכה (ע"ר)
NATIONAL INSTITUTE FOR TESTING & EVALUATION
المركز القطري للامتحانات والتقييم
מיסודן של האוניברסיטאות בישראל

דוח מרכז 388
ISBN:978-965-502-178-3

פיתוח מטלת הכתיבה בתחום המילולי של מכפ"ל

גלעד סופר, ספי פומפיאן, נעמי גפני

מאי 2013

תוכן העניינים

4	רקע
4	מבנה הדוח
6	חלק א – מבוא
6	מדוע צריך להעריך כתיבה
6	סוגיות מרכזיות בפיתוח מטלות כתיבה
15	סקירת מטלות הכתיבה בבחינות מקבילות למכפ"ל
26	חלק ב – פיתוח מטלת הכתיבה במכפ"ל
26	תהליך פיתוח מטלת הכתיבה – החלטות ראשוניות
28	פיתוח דגמי המטלות הבסיסיים
32	ניסוי ראשון
37	ניסוי שני
44	חלק ג – הערכת החיבורים ונקידתם
44	שיטות ההערכה הקיימות
45	סוגי מחוונים
50	תהליך פיתוח המחוון
56	הערכות החיבורים – ניסוי שני
57	הכשרת המעריכים
61	אופן הערכת החיבורים בפועל
63	חלק ד – סוגיות נוספות בפיתוח מטלת הכתיבה במכפ"ל
63	היקף הכתיבה במטלת הכתיבה
64	בחירה
64	הערכה שלישית
66	פרסום המחוון (בחוברת ההסבר ובאתר המרכז)
66	עבודה ללא העברה ניסיונית
67	סיכום
68	ביבליוגרפיה

רקע

פיתוח מטלת הכתיבה התחיל בשלהי 2010, בעקבות החלטת המרכז הארצי לבחינות והערכה (מאל"ו) על סדרת שינויים במבחן הכניסה הפסיכומטרי לאוניברסיטאות (מכפ"ל). בין השאר הוחלט להוסיף מטלת כתיבה לחלק המילולי של הבחינה, ולקצר את פרקי הבררה בבחינה, כך שהזמן המוקצה לפתרונם יתקצר בארבעים דקות (20% מזמן הבחינה). זאת כדי לאפשר את הוספת מטלת הכתיבה, שיוקצו לה 30-35 דקות.

במסגרת פיתוח המטלה, שארך כשנתיים, סקר צוות הפיתוח מבחני כתיבה דומים בעולם, פיתח סוגים חלופיים שונים של מטלות, בלווית מחוונים והליכי הכשרת מעריכים המתאימים להם, וערך שני ניסויים רחבי היקף, שנבדקו בהם חלופות אלה, ושבעקבותיהם הוחלט מה יהיה סוג המטלה ומה יהיו מאפייניה. כמו כן, הצוות החל לבנות בנק מטלות תפעוליות. כל אלה נסקרים בדוח בפרוטרוט.

צוות הפיתוח כלל את חברי התחום המילולי ממדור פיתוח בחינות במאל"ו, בראשותו של ספי פומפיאן. את תהליך הפיתוח ליוו ד"ר דנה טאובה, ד"ר עידית שקד וגב' נחמה פרקר מהיחידה להוראת כתיבה מדעית במכינה הקדם אקדמית ובפקולטה למדעי הרוח באוניברסיטה העברית.

מבנה הדוח

לדוח ארבעה חלקים.

חלק א – המבוא – מציג סוגיות מרכזיות בפיתוח מבחני כתיבה, גישות שונות למדידת יכולת כתיבה, ומטלות כתיבה במבחנים מקבילים למכפ"ל.

חלק ב מציג את תהליך הפיתוח של מטלת הכתיבה במאל"ו ומתאר את ההחלטות הראשוניות שנגעו לפיתוח המטלה, את הדגמים הבסיסיים ששימשו נקודת מוצא לפיתוח, את שני הניסויים שנערכו כדי לבדוק את המטלות וכן את ההחלטות והקווים המנחים שנגזרו מניסויים אלה. תהליך הפיתוח מוצג כרונולוגית ומלווה בדוגמאות.

חלק ג עוסק בהערכת החיבורים ובנקידתם. בחלק זה מוצגות שתי הדרכים העיקריות להערכת כתיבה: הערכה הוליסטית והערכה אנליטית. לאחר מכן מתואר תהליך הפיתוח של המחווה (המדריך להערכת החיבורים) ושל שיטת הנקידה. לבסוף מוצג הליך ההכשרה של מעריכי החיבורים.

חלק ד מוקדש לסוגיות נוספות הקשורות בפיתוח מטלת הכתיבה ובשימוש בה במכפ"ל: היקף הכתיבה, טיפול במקרים של אי-הסכמה בין מעריכים, פרסום המחווה לציבור, הליך העבודה בתחום המילולי ועבודה ללא פיילוט.

החלוקה שלעיל נועדה לפשט את הצגת הליך הפיתוח. בפועל המרכיבים שבכל אחד מהחלקים תלויים זה בזה ופיתוחם נעשה במקביל.

הערה: בשנים 1995-1997 פותח במאלי"ו מבחן כתיבה, בעקבות החלטה של הנהלת המרכז להציע מבחן כזה לאוניברסיטאות כתוספת לבחינה הפסיכומטרית. מכיוון שלא נמצא למבחן קהל יעד מוגדר, ונוכח עלות הפיתוח הגבוהה שלו, החליטה הנהלת המרכז להקפיא את פיתוח המבחן. עם זאת, הידע שנצבר אז שימש את צוות הפיתוח כנקודת מוצא לעבודתו וכן כבסיס לדוח זה, שחלק לא מבוטל מהחומר המוצג במבוא לו נסמך על דוחות צוות הפיתוח מאז: דוח המרכז מס' 214, שחיברה דנה גנאור-שטרן ופורסם בפברואר 1996, והחיבור לסיכום פרויקט מבחן הכתיבה, שכתבו דנה גנאור-שטרן ורותי ריזל ופורסם באוגוסט 1997.

חלק א – מבוא

מדוע צריך להעריך כתיבה

מיומנות הכתיבה ומיומנות הקריאה בלימודים האקדמיים הן יכולות משלימות, ושתייהן הכרחיות להצלחה בלימודים אלה. הקריאה היא הערוץ המרכזי לרכישת ידע; הכתיבה משמשת לסיכום החומר הנלמד, ליישומו ולפיתוח והבעה של רעיונות מקוריים, ובנוסף – למילוי מטלות לימודיות (עבודות ומבחנים), המשמשות בסיס להערכת הישגי הסטודנטים. אף שאין עוררין על חשיבותה של מיומנות הכתיבה בלימודים האקדמיים, בחינות המיון בארץ ובעולם מדגישות יותר את בחינת יכולת הקריאה. הסיבה העיקרית לכך היא שאת יכולת הקריאה והיכולות הנגזרות ממנה קל וטבעי להעריך באמצעות מבחני בררה, ואילו את יכולת הכתיבה קשה להעריך באמצעות מבחנים כאלה. עם זאת, בשנים האחרונות התרחבה מגמת השימוש בהערכות ביצוע המבוססות על מבחנים פתוחים, ובמבחני מיון רבים ללימודים אקדמיים, בעיקר בארצות הברית, שולבו מטלות הבוחנות יכולת כתיבה.

סוגיות מרכזיות בפיתוח מטלות כתיבה

יכולת הכתיבה היא יכולת רב-ממדית המשמשת במגוון הקשרים ומשימות, ותוצריה רבים ומגוונים. מטבע הדברים גם מטלות כתיבה שונות נבדלות אלה מאלה במגוון מאפיינים. הצעד הראשון בפיתוח מטלת הכתיבה היה אפוא בחינה של כמה סוגיות עקרוניות הנוגעות לאפיון מטלות כתיבה: מהימנות, תוקף ואוטנטיביות; בחינה עקיפה לעומת בחינה ישירה; מידת ההתבססות על יכולת קריאה במטלות כתיבה; מידת ההסתמכות על ידע קודם של הנבחנים; תפקיד הגריין (קטע הטקסט שהמטלה מבוססת עליו) במטלה; מידת המיקוד של המטלה. מבחינת הסוגיות האלה נגזרו מאפייניהן של מטלות הכתיבה שפותחו בשלבי העבודה הראשוניים.

מהימנות

מהימנות היא מידת העקיבות במדידתו של דבר-מה. אחד הגורמים המשפיעים ביותר על המהימנות הוא מספר המטלות במבחן. במהימנות עלולים לפגוע גם גורמים אחרים הקשורים למטלות במבחן, וכן גורמים הקשורים לנבחנים וגורמים הקשורים להערכת הבחינה.

גורמים הקשורים למטלה

- **נושא מטלת הכתיבה:** כל נושא לחיבור נטוע בעולם תוכן מסוים, וסביר שנבחן יצליח יותר במטלת כתיבה אם הוא בקי בעולם התוכן שלה. ישנן עדויות מחקריות רבות לקשר חזק בין מידת ההיכרות עם נושא מטלת הכתיבה ובין הצלחה בה (Langer, 1984). במחקר של Breland et al. (1987) נמצא כי האינטראקציה בין נושא הכתיבה ובין הנבחנים מסבירה אחוז גדול מהשונות בהערכת הכתיבה (22%). נראה שרמת הכתיבה מושפעת ממידת קרבתם של הכותבים לנושא והיכרותם עמו, ומכאן שבדיקת

יכולת כתיבה הנסמכת על חיבורים בנושאים שונים מועדת לפגיעה במהימנות, שכן במטלות מוערך גם משתנה מתערב, והוא מידת ההיכרות של הנבחן עם נושא המטלה – היבחנות חוזרת של נבחן במטלה שקולה, אך בנושא אחר, שמוכר לו פחות, עשויה להביא להערכה נמוכה יותר של יכולת הכתיבה שלו. לכן חשוב לא להסתפק במטלה בודדת.

- **סוגת הכתיבה הנדרשת:** במחקרים שונים נמצא כי רמת הכתיבה של נבחנים משתנה לפי הסוגה שהם מתבקשים לכתוב בה. למשל, Crowhurst & Piche (1979) מצאו שרמת התחביר בחיבורים שנכתבו בתגובה למטלות שדרשו כתיבה טיעונית הייתה גבוהה ומורכבת יותר בממוצע מרמת התחביר בחיבורים שנכתבו בתגובה למטלות שדרשו כתיבה תיאורית. Reid (1990) השווה בין המאפיינים הלשוניים של חיבורים שנכתבו בתגובה לשני סוגי מטלות – מטלת השוואה בין טקסטים ומטלת כתיבה טיעונית – ומצא הבדלים ניכרים באורך החיבורים, באוצר המילים בהם ובשימוש בשמות גוף בהם. גם ממצא זה מחזק את השאיפה לכלול בבחינה מטלות כתיבה מגוונות.

גורמים הקשורים לנבחנים

מטלת כתיבה אחת עלולה להיות קשה לנבחנים אחדים יותר מלאחרים מסיבות שאינן קשורות ליכולת הכתיבה שלהם.

- **ידע מוקדם:** כאמור, נמצא קשר מובהק בין מידת ההיכרות המוקדמת עם נושא מטלת הכתיבה ובין ההצלחה בה.
- **רקע תרבותי ולשוני:** מחקרים מראים שהרקע התרבותי של נבחנים הוא מקור נוסף לשונות בציוניהם, מעבר לשונות הרלוונטיות ליכולת המוערכת. כשתבניות לשוניות ורטוריות מתרבות המוצא של נבחן שונות מציפיותיהם של מפתחי הבחינה ומעריכיה, הדבר עלול לפגוע בהערכה שלו.
- **פרשנות המטלה:** קיימות עדויות מחקריות לכך שהדרך שבה נבחנים מפרשים את המטלה משפיעה על הציון שהם מקבלים בה. Murphy ועמיתיו (1982) מצאו הבדלים ניכרים בין פרשנויותיהם של תלמידים למטלות ובין פרשנויותיהם של מורים שהעריכו את חיבוריהם של התלמידים לאותן מטלות עצמן.

גורמים הקשורים להערכת הבחינה

- **מעריכי הבחינה:** מחקרים רבים מעידים כי מאפייניהם של מעריכי בחינות משפיעים על הערכותיהם. בין מאפיינים אלה אפשר למנות את הרקע המקצועי של המעריכים, את הרקע התרבותי והלשוני שלהם ואת הוותק שלהם בהערכה.
- **הכשרת המעריכים:** Freedman (1981) מצא שלטיב הליך ההכשרה שעברו המעריכים השפעה לא מבוטלת על השונות הלא רלוונטית בציוני מטלות כתיבה. מהימנות ההערכות של מעריכים שזכו להכשרה מעמיקה שכללה הצגת רקע מפורט של דרך

ההערכה הנדרשת הייתה גבוהה מזו של מעריכים שהוצגו בפניהם רק נושא המטלה ומשמעותו.

- **גורמים קונטקסטואליים** : נמצא כי סדר ההערכה של חיבורים משפיע על הערכותיהם : חיבור נתון יקבל ציון גבוה יותר ככל שהחיבור שקדם לו היה חלש יותר, ולהפך (Sweedler-Brown, 1993).

תוקף

תוקף המבחן הוא המידה שבה ההחלטות המתקבלות על סמך ציוני המבחן, או הפרשנות שניתנת להם, מוצדקות על פי עדויות ממקורות שונים. יש עדויות וממצאים מסוגים שונים המבססים את תוקף המבחן ומצדיקים את השימוש בציונים למטרה שהמבחן נבנה לשמה :

תוקף תוכן

תוקף תוכן הוא המידה שבה אפשר להסיק מציון במבחן על התכונה שהמבחן מבקש להעריך. כדי לבסס את תוקף התוכן של מבחן כתיבה יש לנתח את יכולת הכתיבה, לגבש הגדרה ממצה של מרכיביה ולבנות את המבחן כך שייצג את כל המרכיבים הרלוונטיים – ורק אותם – במשקל המתאים לחשיבותם היחסית. מחקרים על מבחני כתיבה ישירים הראו כי יש שונות בהערכות יכולת הכתיבה של אותו אדם כשהן ניתנות למטלות כתיבה מסוגים שונים. לכן כדי להגדיל את תוקף התוכן של מבחן כתיבה, יש להתמקד רק במיומנויות הכתיבה שמבקשים למדוד.

תוקף ניבוי

תוקף הניבוי בודק את ביצועי הנבחן מול קריטריון : מדד ישיר ובלתי תלוי של היכולת שהמבחן נועד למדוד. נשאלת השאלה מהו הקריטריון הרלוונטי לבדיקת התוקף של מטלות כתיבה. בארצות הברית נבדק תוקף הניבוי של פרק הכתיבה של מבחן ה-SAT כנגד שני קריטריונים מרכזיים, הנבדלים במידת הקרבה שלהם לחזאי (מבחן הכתיבה) – ציוני הקורסים באנגלית הכלולים בלימודי התואר הראשון, וממוצע ציוני שנה א בכלל הקורסים בלימודים לקראת התואר.

תוקף מבנה

תוקף מבנה הוא המידה שבה המבחן מודד מבנה או תכונה תיאורטיים. אי אפשר לצפות במבנים כאלה ישירות, ולכן הם מתורגמים לעדויות התנהגותיות. מבנה מוגדר בשתי רמות : הראשונה – הגדרה אופרציונלית, כלומר תרגום המבנה למדדים שאפשר לצפות בהם, והשנייה – העלאת השערות הנוגעות לקשרים בין המבנה התיאורטי ובין מבנים אחרים ובדיקה אמפירית שלהן.

תוקף השלכתי

Messick (1989) טוען כי בעת הערכת התוקף של מבחן יש להביא בחשבון את ההשלכות החברתיות – הן ההשלכות הפוטנציאליות, הן ההשלכות בפועל – של השימוש בו. התוקף ההשלכתי של מבחן גדל ככל שההשלכות של השימוש במבחן חיוביות יותר. סביר להניח כי

הוספת מטלת כתיבה למבחן הכניסה למערכת ההשכלה הגבוהה תגדיל את ההשקעה בשיפור יכולת הכתיבה בקרב המבקשים ללמוד בה.

אותנטיות

מושג מרכזי נוסף בהערכת התוקף של מטלות כתיבה הוא האותנטיות – המידה שבה המטלה מייצגת את הכתיבה שרוצים להעריך בפועל.

יש כמה סוגי מטלות המשמשים להערכה ישירה של יכולת הכתיבה :

כתיבה תיאורית

למשל, תיאור תמונה נתונה. חסרונותיו הבולטים של דגם זה כאמצעי להערכת יכולת כתיבה אקדמית הם היעדר אותנטיות וכן שטחיות הדגימה של יכולת הכתיבה.

חיבור פתוח

כתיבה על נושא נתון, אולם ללא חומר עזר או עיגון בהקשר כלשהו. גם דגם זה נעדר אותנטיות בהקשר של הערכת יכולת כתיבה אקדמית. בנוסף, הוא סובל במיוחד מתלות בידע בתחום הדעת, ומהיעדר קריטריונים ברורים מספיק לכתיבה, מה שעלול לגרום לקשיים בשלב ההערכה.

חיבור מובנה

כתיבה על נושא נתון המעוגנת בסיטואציה חברתית ומכוונת לנמען מסוים. דגם זה אותנטי יותר וממוקד יותר, ולכן מאפשר הערכה מהימנה יותר. ואולם גם הוא מבוסס במידה רבה על ידע ודורש מהנבחן לדמות לעצמו את הסיטואציה החברתית המוגדרת במטלה.

חיבור תגובה לגריין נתון

הנבחן מקבל גריין (בדרך כלל הגריין הוא טקסט, אך לעתים גם נתונים בפורמט של טבלאות, גרפים, גריונים אודיו-ויזואליים וכדומה משמשים גריינים), ועליו לכתוב חיבור המתייחס לגריין זה. המטלות הנדרשות עשויות להיות מגוונות :

- סיכום הקטע או הנתונים (זו למשל הדרישה במבחן TOEFL)
 - ניתוח הטקסט או הנתונים אגב הפגנת הבנה של המידע המוצג בהם
 - התייחסות ביקורתית-אנליטית לקטע, לנאמר בו ולמבנהו
 - נקיטת עמדה מנומקת בנוגע למידע המוצג בקטע
- מרבית המבחנים שיוצגו בהמשך מבוססים על מטלות חיבור משני הסוגים האחרונים, והם בודקים יכולות כתיבה גבוהות יותר מאלה שנבדקות בשני הסוגים שלפניהם, שכן היכולות שנדרשות בסוגים הראשונים הן תנאי הכרחי, אך לא מספיק, להצלחה באחרונים.

בניית מאמר

הנבחן נדרש לבנות מאמר חדש על סמך שילוב של רעיונות מתוך כמה מקורות הניתנים לו בבחינה.

מבין הדגמים לעיל, לשני האחרונים אותנטיות גבוהה כמטלות מיון הבודקות יכולת כתיבה עיונית, הכתיבה האופיינית ללימודים באקדמיה. יתרון נוסף שלהם נעוץ בכך שהם אינם נשענים במידה רבה על ידע בתחום הדעת. עם זאת, יש לזכור כי ככל שהגריין שהמטלה מתבססת עליו ארוך ועשיר יותר, רמת הביצוע בה תהיה תלויה יותר ביכולת הבנת הנקרא, שממילא נבדקת בשאלות הבררה בבחינה. בנוסף, גרייני המטלות מדגמים אלה – ובמיוחד מדגם בניית המאמר – דורשים זמן ארוך יחסית לקריאתם.

בחינה עקיפה מול ישירה

אף שכל מבחן כתיבה הוא למעשה אומדן עקיף של יכולתו של הנבחן, המושגים "בחינה עקיפה" ו"בחינה ישירה" משמשים להבחנה בין שתי גישות בסיסיות בהערכת כתיבה:

מבחני כתיבה עקיפים

מבחני כתיבה עקיפים משמשים להערכה של יכולת הכתיבה באמצעות תצפיות על סוגים של ידע ומיומנויות הקשורים לכתיבה. הפריטים במבחנים אלה הם פריטי בררה, שאינם דורשים יצירה אקטיבית של טקסט, אלא זיהוי דוגמאות של כתיבה תקינה או של טעויות כתיבה. מטבע הדברים, פריטים אלה דורשים לפתרונם זמן קצר, ולכן מבחנים עקיפים כוללים פריטים רבים יחסית. סוגי המטלות הנפוצים במבחני כתיבה עקיפים הם תיקון משפט, תיקון פסקה וזיהוי טעויות במשפט (דוגמאות יובאו בהמשך).

מבחני כתיבה ישירים

מבחני כתיבה ישירים כוללים מטלות כתיבה אקטיבית של חיבורים או מאמרים. הנבחנים מתבקשים לחבר טקסט (אחד או יותר), בדרך כלל בזמן קצוב, על נושא או נושאים נתונים. בגלל אורך הזמן הנדרש לכתיבת חיבור, מבחני הכתיבה כוללים בדרך כלל מעט משימות כתיבה. מבחני הכתיבה הישירים מדגישים את היות הכתיבה תהליך יצרני ואקטיבי שניתן בו ביטוי ישיר ואופטימלי לתהליכים הקוגניטיביים ולהיבטים הפונקציונליים והתקשורתיים שביכולת הכתיבה של הנבחנים.

מטלות כתיבה עקיפות הן אובייקטיביות, אפשר לכלול רבות מהן בבחינה מוגבלת בזמן, והן מאפשרות כיוול בין נוסחים. אך הן מתייחסות רק לחלקים קצרים מתוך טקסט, ואינן מודדות באופן ישיר יצירה של טקסט, אלא רק טווח מוגבל של יכולות הקשורות לכתיבה, כגון זיהוי טעויות דקדוק ותחביר. לפיכך, למטלות מסוג זה אותנטיות נמוכה. לעומתן, למטלות ישירות אותנטיות גבוהה: הן מעריכות את יכולת הכתיבה של הנבחנים על ידי הערכת תוצרי הכתיבה שלהם, והערכה זו מתייחסת ליכולות גבוהות – כשרים הקשורים לרמה הכוללת והגבוהה ביותר של כתיבת טקסט. ואולם, גם למטלות הישירות תוקף נראה מוגבל: סיטואציית המבחן ומאפייניה (מגבלות הזמן, הלחץ, אקראיות הנושא ועוד) בודקים את היכולת לייצר בלחץ זמן טיוטה מהירה על נושא נתון, מבלי לתת ביטוי לשלבים מהותיים של תהליך הכתיבה האקדמית, כמו עריכה, תיקונים, מחקר ואינטגרציה של נתונים.

למטלות הכתיבה הישירות עוד כמה חסרונות: ראשית, תהליך הנקידה מבוסס על הערכות סובייקטיביות ומכניס רעש למדידה. אמנם בבדיקות שנעשו בארצות הברית עם הכנסת מבחן הכתיבה ל-SAT נמצאה רמה גבוהה יחסית של מהימנות בין שופטים (0.81-0.77) (Proctor, T. P., 2009, & Young Kounig, R.), אך עדיין מדובר בפגיעה במהימנות לעומת זו של מבחנים עם הערכה אובייקטיבית. שנית, ההערכות הן בדרך כלל הוליסטיות (או הוליסטיות למחצה – גם כשהן מוגדרות אנליטיות), ולכן מועדות לאי-דיוקים. שלישית, סולמות הנקידה בדרך כלל מוגבלים (במרבית מבחני הקבלה משתמשים בסולם 1-6), דבר שגורם להשטחת הבדלים בין נבחנים. רביעית, מהימנות המבחנים נמוכה יחסית למבחנים עקיפים בגלל מיעוט הפריטים. ולבסוף, ביצועי הנבחנים תלויים במידת-מה בתוכן (תחום הדעת) ובסוגת הכתיבה הנדרשת, ומבחן שכולל משימות כתיבה מועטות יכול להעריך רק את היכולת בסוגות ובתכנים שנכללים בו – ומכאן גם הקושי להעריך את רמת הקושי של המטלות השונות ולכילל בין נוסחים.

מידת ההתבססות על קריאה במטלות כתיבה

גישות מחקריות שונות מפנות את תשומת הלב לקשר ההדוק בין יכולת הקריאה ויכולת הכתיבה. עם זאת, אחד המשתנים הבולטים המבדילים בין מטלות כתיבה שונות הוא מידת ההסתמכות שלהן על קריאה. מידת ההסתמכות של המטלה על קריאה תלויה בשלושה גורמים: אורך הטקסט של הגריין ושל המטלה, רמתם הלשונית והמידה שבה הנבחנים נדרשים להסתמך בכתיבתם על הכתוב בגריין. יש שונות ניכרת בשלושת המשתנים הללו במטלות הכתיבה שיוצגו בסקירת המבחנים להלן.

קריאה-כתיבה – הציר הקוגניטיבי

בשנות השבעים והשמונים של המאה העשרים רווחו שתי גישות בולטות שניסו לעמוד על טיב הקשר בין קריאה לכתיבה: הגישה המתערבת והגישה המתאמת. הגישה המתערבת התמקדה בבדיקת ההשפעה שיש לתרגולה של יכולת אחת על הביצועים ביכולת האחרת. מרבית המחקרים בגישה זו מצאו שאכן תרגול של כל אחת מהיכולות מביא לשיפור ביכולת האחרת.

נמצא למשל שיכולת הארגון בכתיבת טקסט השתפרה בעקבות לימוד שהתמקד בקריאת טקסטים אגב הדגשת מאפייניהם המבניים. הגישה המתאמת מדדה את הקשרים הסטטיסטיים בין מדדי יכולת קריאה וכתיבה, ומצאה קשרים ניכרים ביניהם. Shanahan (1984) מצא שהשונות הנצפית בכל אחת מהיכולות יכולה להסביר כמעט מחצית מהשונות ביכולת האחרת, ובניתוחים שבחנו את כיוון הקשר בין שתי היכולות העלו ניתוחיו שההשפעה היא דו-כיוונית.

המודל הקוגניטיבי המקובל (Flower & Hayes, 1981) רואה את תהליך הכתיבה כמורכב משלושה תהליכי משנה: תכנון (שמורכב מיצירת רעיונות,



ארגונים והצבת מטרות לכתיבה), יצירת הטקסט (תרגום המידע למשפטים קבילים בשפה) וביקורת (קריאת הטקסט המופק ועריכתו). לביקורת, הכרוכה בקריאת הטקסט הכתוב, שני תפקידים עיקריים בתהליך הכתיבה: היא גוררת כתיבה של טקסט נוסף, בהתאם למה שנקרא, ומביאה גם לעריכת תיקונים בטקסט שכבר נכתב. בסיטואציה של מבחן כתיבה (ר' סרטוט) מתווספת למודל גם "סביבת המבחן" (מטלת הכתיבה, שמציגה את נושא הכתיבה וכוללת גם מידע נוסף, והטקסט שנכתב עד כה), ומתקיים היזון חוזר בינה ובין תהליך הכתיבה. בתהליך הכתיבה, כפי שהוא מוצג במודל, מעורבת יכולת הקריאה במידה ניכרת בשני תהליכי משנה: ראשית, קריאת המידע המופיע בגריין ובמטלה היא הבסיס לתכנון הטקסט. שנית, תהליך הביקורת מבוסס על קריאת הטקסט שנכתב.

קריאה-כתיבה – הציר הפונקציונלי

הגישה הסוציו-לינגוויסטית רואה בקריאה ובכתיבה פעולות משלימות הנטועות במציאות חברתית מוגדרת. אף שהכתיבה נעשית בדרך כלל בידי אדם אחד, הרי שהיא נעשית בסיטואציה חברתית, שכן הכותב מייעד את דבריו לקהל מוגדר וכותב אותם למען מטרה מסוימת. לפי גישה זו, גם את משימות הכתיבה יש לקחת מהקשר חברתי מוגדר (Bloom & Green, 1984). גישות אורייניות מודרניות גורסות כי יש לראות בקריאה ובכתיבה מרכיבים שונים של אותה מיומנות. De Beaugrande (1980) טוען כי בכתיבה ובקריאה מתקיימות אותן אופרציות, אך בסדר הפוך: הכותב מתרגם את כוונותיו לתוכן מסוים ומנסח תוכן זה בטקסט כתוב, ואילו הקורא גוזר מהטקסט הכתוב את התוכן וממנו מקיש על כוונת הכותב (ההגעה של הקורא לכוונותיו המקוריות של הכותב תלויה כמובן ביכולות הקריאה והכתיבה של השניים, בהתאמה).

ההפרדה בין קריאה לכתיבה מטושטשת גם מפני ששתי הפעולות נעשות בדרך כלל במסגרת אותה משימה, הן כאשר הכתיבה נסמכת על קריאת מקורות, כפי שקורה בדרך כלל בעולם האקדמי, והן כאשר קריאה כזאת אינה הכרחית, שכן גם אז, כאמור, הכותב קורא את הטקסט שכתב בעת ביקורת הטקסט ועריכתו.

Sarig & Folman (1988) טוענות כי הקשר ההדוק בין תהליכי הכתיבה והקריאה נובע משלושה גורמים: ראשית, בין תהליכי הקריאה והכתיבה יש אינטראקציה, ומתנהל ביניהם משא ומתן על משמעות הטקסט; שנית, בשני התהליכים מתרחשות פעולות קוגניטיביות מקבילות בסדר הפוך; ושלישית, בהקשר האקדמי שתי הפעולות מתבצעות בדרך כלל יחד, כך שבכל משימה הדורשת כתיבה נדרשת גם קריאה. לכן, לדבריהן, אין ללמד קריאה וכתיבה בנפרד ואין לבחון אותן בנפרד.

מידת ההסתמכות על ידע קודם במהלך הכתיבה

עדויות מחקריות רבות מראות קשר מובהק בין מידת ההיכרות עם נושא הכתיבה ובין הצלחה במטלה (ראו הדיון במהימנות לעיל). כדי שהמבחן יהיה הוגן, מהימן ותקף ככל האפשר, יש לצמצם אפוא את ההסתמכות על ידע קודם בתחומי הדעת הנדונים בבחינה. בדרך כלל, מבחן כתיבה מכוון את הנבחן לכתיבה בנושא מסוים, ולכן הוא מודד לא רק את כושר הכתיבה שלו אלא גם את התמצאותו בעולם התוכן המסוים שהחיבור יעסוק בו.

אפשר לסווג מטלות כתיבה לארבע קבוצות, לפי מידת ההסתמכות על ידע קודם :

1. מטלה הבודקת יכולת הבעה בכתב של ידע ספציפי בעולם תוכן נתון. משימה זו מתאימה למבחני הישג, והיא המטלה השכיחה ביותר בבתי הספר ובאוניברסיטאות.
2. מטלה המתייחסת לתחום נתון באופן כללי, ומאפשרת לנבחן לבחור באיזה ידע ישתמש לכתיבת החיבור. במטלה כזאת יש הסתמכות על ידע בתחום דעת מסוים, אך ללא הגדרה של ידע ספציפי שנדרש.
3. ביסוס משימת הכתיבה על מידע נתון. בראשית המבחן ניתן לנבחן מידע (קטע קריאה, תרשים או טבלה) ועל הנבחן לכתוב חיבור המתייחס למידע שניתן לו. המטלה יכולה להיות מסוגים שונים – סיכום הנאמר בטקסט, הבעת עמדה בנוגע לכתוב, יישום הטקסט על תחומים אחרים ועוד. לחלופין אפשר לכלול כמה קטעים קצרים, ולדרוש מהנבחן ליצור אינטגרציה בין הרעיונות השונים המובאים בקטעים, או להעריך אותם. יתרונה של שיטה זו הוא צמצום ההשפעה של היכרות עם עולם תוכן מסוים על הציון שיקבל הנבחן במבחן הכתיבה. בנוסף, היא דומה יותר למטלות שסטודנטים נדרשים לבצע באוניברסיטה – קריאת טקסט, סיכום והערכתו. מאפיין חשוב של השיטה, ויש שיאמרו – חיסרון בולט שלה, הוא היעדר הבחנה ברורה בין מדידת הבנת הנקרא למדידת יכולת הכתיבה, שכן הבנת הקטע היא תנאי מקדים להצלחה בכתיבה.
4. מטלה המבוססת על גריין כללי ומינימלי. משימה זו כללית יותר מהאחרות, אינה מתמקדת בתחום מסוים ומאפשרת לנבחן לבחור הן את הידע הספציפי והן את התחום הכללי שיכתוב עליו. ב-GRE וב-SAT יש שימוש במטלות כאלה. יתרונה הגדול של סוג מטלות זה הוא הקטנת השפעתו של הידע המוקדם על הביצוע. חסרונה העיקרי הוא המרחב הגדול שהוא מותיר לפרשנות – בין הנבחנים עשויים להיות הבדלים ניכרים באופן הבנת השאלה, ומטבע הדברים יהיו הבדלים בנושאי הכתיבה. כך ייווצרו למעשה כמה נוסחים של אותה בחינה, והדבר יקשה על הערכה מהימנה של החיבורים.

תפקיד הגריין במטלה

בין הגריין ובין המטלה יכולים להתקיים קשרים מסוגים שונים :

- מטלה המבוססת על הבנת הגריין – מטלה שאי אפשר לבצעה ללא הבנה של המידע המוצג בגריין. לדוגמה :

הקטע הבא נלקח ממאמר מאת ריקי שיו, שהופיע בכתב העת "מעשי משפט". קראו אותו בעיון וענו על מטלת הכתיבה שלאחריו :

קצבת הבטחת הכנסה היא קצבה שמטרתה להגן על מי שהכנסותיו אינן מספיקות כדי להבטיח לו רמת חיים מינימלית. בשנת 2003 הנהיג שר האוצר בארץ תכנית חירום כלכלית שכללה, בין השאר, קיצוץ ניכר בקצבת הבטחת ההכנסה. כמה ארגונים חברתיים טענו כי קיצוץ זה פוגע פגיעה קשה בכבודה של אוכלוסיית מקבלי הבטחת הכנסה בישראל, שהיא האוכלוסייה החלשה ביותר מבחינה כלכלית. מנגד, טענו נציגי הממשלה כי רק "היעדר קורת-גג, רעב והיעדר כסות, יחשבו כפגיעה המחללת את כבודו של אדם".

האם אתם מסכימים עם עמדת הממשלה באשר לקיום בכבוד? נמקו את דבריכם.

- מטלה שהגריין מספק בה הקשר בלבד, ואינו הכרחי לביצועה. לדוגמה:

מגבלותיו של המחשב ככלי עזר לרופא – מבוסס על מאמר מאת ד"ר מירון אפשטיין

משעה שנכנס המחשב לרפואה, חלומם של רופאים רבים היה להגיע למצב שבו יוכלו להזין את המחשב בנתונים על חולה מסוים, ולקבל ממנו אבחנה מוגמרת. כיום מצויות תוכנות אבחון המסוגלות לספק לרופאים רשימה של אבחנות אפשריות המתאימות לאוסף נתונים מסוים.

קיימות מספר בעיות אשר מונעות מהרופאים להסתמך באבחנותיהם באופן בלעדי על המחשב. המחשב יכול להתחשב רק בנתונים בעלי אופי כמותי, כמו טמפרטורת הגוף. אולם קיימים נתונים רפואיים רבים שאינם ניתנים לכימות, למשל צורה של פריחה בעור. בנוסף, מידת החשיבות של נתונים רפואיים שונים בקביעת האבחנה אינה קבועה, אלא משתנה על פי ההקשר. כך למשל, הרופא עשוי להעניק חשיבות רבה בהקשר מסוים לטמפרטורת גוף חורגת מהנורמה, ואילו בהקשר קליני אחר לא לייחס משקל רב לאותה טמפרטורה.

בעיה נוספת היא שהרופאים מנסים בדרך כלל למצוא אבחנה אחת שתסביר את כל הסימנים של החולה. אולם לעיתים נמצאים סימנים החורגים מהנורמה אך אינם מעידים על מחלה מסוימת. סימנים אלה הם בבחינת "רעש" שעלול להסיח את דעתו של הרופא מן האבחנה המדויקת. ייתכנו גם מצבים שבהם יופיעו סימנים שאינם מתיישרים לדעתו של הרופא עם האבחנה המקורית, אך הם אינם מקריים, אלא מעידים על מחלה נוספת שממנה סובל החולה. היכולת להבחין בין שני המצבים הללו נעשית באמצעות כלים שיש בבעלות הרופא ואי אפשר להקנותם למחשב, כמו אינטואיציה. מסיבות אלה לא זכו תוכנות האבחון להצלחה גורפת. הן נותרו מוגבלות מאוד, ומשמשות רק לקבוצות מוגדרות היטב של נתונים. בתחומים אלה יכול היה המחשב לספק הנחיות להמשך הברור האבחנתי, ולפעמים גם לעדכן את הרופא בנתונים לגבי סבירות האבחנות האפשריות השונות.

לאחר ביצוע האבחנה, לעתים קרובות יש להתאים טיפול לחולה. גם בתחום זה נעשה ניסיון להיעזר בכלים ממוחשבים. החיסרון של כלים אלה, שוב, נובע מאי-היכולת לכמת את כל הנתונים החשובים הקשורים לחולה מסוים. מסיבה זו כל ההמלצות הטיפוליות שיכול המחשב לספק נוגעות לחולה ה"אידיאלי", כלומר, לחולה דמיוני ומופשט, המייצג את התכונות הטיפוסיות של קבוצת החולים במחלה מסוימת אחת. במקרה כזה יהיה הטיפול אחיד לכל מחלה. אלא שבמצב כזה, מי בכלל זקוק למחשב כמקבל החלטות?

חברת כוח אדם העוסקת במיון מועמדים למשרות שונות בחברת תעופה על פי נתונייהם האישיים (כמו למשל השכלה, ניסיון בעבודה, יצירתיות, יחסי אנוש) שוקלת לפתח תוכנת מחשב שתבצע את המיון. התקווה היא כי תוכנה זו תוכל להחליף בעתיד את המומחה המבצע כיום עבודה זו. דונו באפשרות להשתמש בתוכנה זו לאור הקטע שקראתם.

מחד גיסא, מטלות הנסמכות במידה רבה על הבנת הגריין – כלומר, על יכולת הבנת הנקרא – הן אותנטיות יותר, אך עלולות להקטין את השונויות בציונים הנובעת מיכולת הכתיבה. מאידך גיסא, שימוש בגריין רק כהקשר מביא לבזבוז זמנו של הנבחן ולרמה נמוכה של אותנטיות של המטלה.

מידת המיקוד של המטלה

ככל שהמטלה ממוקדת יותר, כך היא מאפשרת השוואה קלה וסטנדרטית יותר בין החיבורים הנכתבים בתגובה לה. עם זאת, מטלות ממוקדות עלולות להגביל את אפשרויות הכתיבה של הנבחנים לתחום צר, שאינו נוח לכולם במידה שווה (בשל הבדלים בידע המוקדם בתחום, בנטיית אישיות וכדומה).

■ דוגמה למטלה ממוקדת

בעבר במרבית מדינות העולם ניתנה הזכות להצביע בבחירות לפרלמנט בגיל 21, אך זה כמה עשורים ניכרת מגמה עולמית להוריד את גיל הבחירה ל-18. בכמה מדינות, ובהן ישראל, הוגשו הצעות חוק להורדת גיל הבחירה מ-18 ל-17 ואף ל-16, ויש מדינות שכבר עשו כן. בדיון הציבורי על הורדת גיל הבחירה נשקלות שאלות כגון: מאיזה גיל אדם זכאי להשפיע על הרכב הפרלמנט, המקבל החלטות שמשפיעות השפעה ישירה על חייו? האם בני הנוער בשלים דיים להחליט החלטות המשפיעות על הציבור כולו? מה יהיו ההשלכות של הכללת בני הנוער בקבוצת הבחורים על המעורבות של בני הנוער בנעשה במדינה ועל המערכת הפוליטית בכללה?

לדעתכם, האם יש להוריד את גיל הבחירה לכנסת? נמקו. תוכלו להיעזר בשאלות המוצגות בפסקה.

■ דוגמה למטלה לא ממוקדת

קיומו של רצון חופשי אינו אלא אשליה – גורלנו נקבע על פי הנסיבות שאליהן נולדנו.

הביעו דעתכם על טענה זו.

סקירת מטלות הכתיבה בבחינות מקבילות למכפ"ל

כשלב מקדים לפיתוח מטלת הכתיבה נסקרו מטלות הערכת הכתיבה במבחנים מקבילים למכפ"ל, בעיקר בארצות הברית, וכן מבחן הכתיבה שפותח במאל"ו בשלהי שנות התשעים ומבחני הבגרות בהבעה. כל הבחינות שיוצגו להלן משתמשות בפריטי כתיבה ישירים: בכולן הנבחנים נדרשים לכתוב סוגים שונים של חיבורים בעצמם. שתי הבחינות האמריקניות לקראת תואר ראשון כוללות, בנוסף לכך, גם פריטי כתיבה עקיפים. בשום מקרה אין הערכה של כתיבה באמצעות פריטים עקיפים בלבד, ואף לא באמצעות פריטים ישירים שאינם חיבור.

בחינות מיון בארצות הברית

בראשית דרכה אופיינה מערכת המיון האמריקנית בשימוש של המוסדות השונים במטלות כתיבה ישירות. למשל, בשלהי המאה התשע עשרה נדרשו מועמדים לאוניברסיטת הרווארד לכתוב סדרת חיבורים המתייחסים ליצירות ספרות מתוך רשימה שהוצגה להם מראש. בראשית שנות השלושים של המאה העשרים ערך ועד ראשי המוסדות (The College Entrance Examination Board) סדרה של שינויים בשיטות ההערכה, כדי להעלות את מהימנות הבחינות ולהוזיל את עלויות הבדיקה שלהן, וכן כדי להפחית את השפעת הרקע הסוציו-אקונומי של הנבחנים על ביצועיהם בבחינות, במטרה לאפשר מוביליות חברתית גדולה יותר. שינויים אלה התבססו על מעבר לשימוש נרחב בפריטי בררה להערכת הכתיבה. במשך שנים רבות הסתמך מרכיב הערכת הכתיבה במבחני המיון החשובים להשכלה הגבוהה בארצות הברית על הערכה עקיפה של יכולת הכתיבה, אך בשני העשורים האחרונים נוספו מטלות כתיבה למרבית המבחנים האלה, מתוך הכרה בכך שמבחן כתיבה תקף חייב לכלול מטלת כתיבה ישירה. המבחן הראשון שהחליף את פריטי הבררה העקיפים במטלת כתיבה ישירה היה ה-LSAT (בחינת מיון ללימודי

משפטים) בשנת 1982, ואחריו עברו מרבית בחינות הקבלה ללימודים אקדמיים שינוי שכלל הוספה של מטלת כתיבה ישירה (או כמה מטלות), לצד הפריטים העקיפים או במקומם.

הבחינות שיוצגו בסקירה להלן הן: SAT (מיון לתואר ראשון), ACT (מיון לתואר ראשון), GMAT (מנהל עסקים – תארים מתקדמים), GRE (תארים מתקדמים), LSAT (משפטים) ו-MCAT (רפואה).

סוגים של פריטי כתיבה עקיפים

פריטי כתיבה עקיפים משמשים בשתי בחינות המיון לתואר ראשון, ה-SAT וה-ACT.

דוגמאות לסוגי הפריטים שבשימוש ב-SAT:

▪ תיקון משפט

הוראות:

קראו את המשפט וקבעו אם החלק המסומן בקו תחתון נכון או שיש לתקנו.
לאחר מכן קראו את אפשרויות התשובה, המחליפות את חלק המשפט המסומן בקו תחתון, וקבעו איזו מאפשרויות השיבוץ תיצור משפט בהיר, מדויק ותקין לשוני.

Laura Ingalls Wilder published her first book and she was sixty-five years old then.

- (A) and she was sixty-five years old then
- (B) when she was sixty-five
- (C) at age sixty-five years old
- (D) upon the reaching of sixty-five years
- (E) at the time when she was sixty five

▪ תיקון פסקה

הוראות:

בחרו באפשרות הטובה ביותר מבין אפשרויות התשובה, גם אם אתם חושבים שעשויה להיות תשובה נכונה נוספת.

- (1) Many times art history courses focus on the great "masters," ignoring those women who should have achieved fame. (2) Often women artists like Mary Cassatt have worked in the shadows of their male contemporaries. (3) They have rarely received much attention during their lifetimes.
- (4) My art teacher has tried to make up for it by teaching us about women artists and their work. (5) Recently she came to class very excited; she had just read about a little-

known artist named Annie Johnson, a high school teacher who had lived all of her life in New Haven, Connecticut. (6) Johnson never sold a painting, and her obituary in 1937 did not even mention her many paintings. (7) Thanks to Bruce Blanchard, a Connecticut businessman who bought some of her watercolors at an estate sale. (8) Johnson is finally starting to get the attention that she deserved more than one hundred years ago. (9) Blanchard now owns a private collection of hundreds of Johnson's works — watercolors, charcoal sketches, and pen-and-ink drawings. (10) There are portraits and there are landscapes. (11) The thing that makes her work stand out are the portraits. (12) My teacher described them as "unsentimental." (13) They do not idealize characters. (14) Characters are presented almost photographically. (15) Many of the people in the pictures had an isolated, haunted look. (16) My teacher said that isolation symbolizes Johnson's life as an artist.

In context, which is the best revision to the underlined portion of sentence 3 (reproduced below)?

They have rarely received much attention during their lifetimes.

- (A) In fact, they had
- (B) Too bad these artists have
- (C) As a result, these women have
- (D) In spite of this, women artists
- (E) Often it is the case that the former have

■ זיהוי טעויות במשפט

הוראות:

קראו את המשפט שלפניכם, ובחנו את החלקים D-A המסומנים בקווים תחתונים.

בחרו את החלק המסומן בקו תחתון שיש לשנותו כדי להפוך את המשפט לתקין.

The other delegates and him immediately accepted the resolution drafted by the neutral

A **B** **C** **D**

states. No error.

E

פריטי כתיבה ישירים משמשים, כאמור, בכל הבחינות שנסקרו, והם נבדלים זה מזה באופי הגריין (קטע הטקסט שהמטלה מבוססת עליו), באורכו ובמורכבותו, באופי הדרישות של מטלת הכתיבה ובמספר מטלות הכתיבה שבכל בחינה ובחינה.

SAT (תואר ראשון)

מבחן הכתיבה מורכב משלושה סוגים של פריטי בררה, הבודקים תפיסת עקבות של זמנים ועקבות דקדוקית, הבנה של תהליך הצגת רעיונות באופן הגיוני, הימנעות משימוש בשפה מעורפלת ודו-משמעית והבנת קואורדינציה וסובאורדינציה (מבנה משפט). מפרט המבחן העקיף, שמוקצות לפתרונו 35 דקות הוא :

1. תיקון משפטים (25 פריטי בררה)
 2. זיהוי טעויות במשפטים (18 פריטי בררה)
 3. שיפור פסקאות (6 פריטי בררה)
- בחלקו השני של המבחן מופיעה מטלת כתיבה ישירה, ולביצועה מוקצות 25 דקות. בחלק זה מוצג לנבחנים קטע קצר והם מתבקשים לכתוב חיבור המציג עמדה מנומקת בנוגע לקטע וראיות שתומכות בעמדה זו. מטלת הכתיבה הישירה נועדה להעריך את יכולת הנבחן להשתמש בשפה מדויקת ונכונה, להציג רעיונות באופן הגיוני, לפתח נקודת מבט ולבטא רעיונות בבירור. משקל בציון : פריטים עקיפים - 70%, חיבור - 30%.

▪ דוגמאות לנושאי חיבור הלקוחים מבחינת SAT

1

אנחנו מוכוונים מאוד לאינדיבידואליזם. אנחנו רואים הכול במונחים של עצמאות אישית, הנאה אישית, הגשמה עצמית. כמו שאומרים: "עשו דברים בשביל עצמכם". הרעיון שאנשים יכולים ממש לעשות דברים בשביל משהו או משהו אחר – קהילה, בית ספר או כל קבוצה אחרת – אבד. עם זאת, חשוב להכיר בכך שכל האנשים מחוברים זה לזה. איננו יכולים לשרוד זה בלעדי זה.

[מבוסס על דבריו של וילארד גיילין בספרו של ביל מוירז "עולם של רעיונות"]

האם אנשים מדגישים מדי עשיית דברים בעצמם ובשביל עצמם? תכננו וכתבו חיבור ובו תסבירו את נקודת מבטכם בסוגיה זו. הביאו תימוכין לעמדתכם באמצעות טיעונים הגיוניים ודוגמאות הלקוחות מתוך טקסטים שקראתם, לימודיכם והתנסויות אישיות או התרחשויות שהייתם עדים להן.

2

רבים מאמינים שכדי לטפס בסולם ההצלחה וההישגים עליהם לשכוח את העבר, להדחיק אותו ולוותר עליו. אולם יש המאמינים בעמדה ההפוכה בדיוק. הם רואים בזיכרונות ישנים הזדמנות להתמודד עם העבר ולשלב אותו בהווה.

[מבוסס על דבריה של שרה לורנס-לייטפוט בספרה "ידעתי גם נהרות : חיים של אבדן וחירות"]

האם זיכרונות מפריעים לאנשים במאמץ שלהם ללמוד מהעבר ולהצליח בהווה, או עוזרים להם? תכננו וכתבו חיבור ובו תסבירו את נקודת מבטכם בסוגיה זו. הביאו תימוכין לעמדתכם באמצעות טיעונים הגיוניים ודוגמאות מתוך טקסטים שקראתם, לימודיכם והתנסויות אישיות או התרחשויות שהייתם עדים להן.

נקידה – שני מערכים בלתי תלויים נותנים כל אחד ציון בסולם 1-6. סכום הציונים הוא ציון הגלם של הנבחן. פער של יותר מנקודה אחת מוביל להערכה שלישית.

ACT (תואר ראשון)

1. פריטי בררה (זיהוי טעויות במשפט, שיפור משפטים ושיפור פסקאות)
 2. מטלת כתיבה ישירה: על הנבחן להציג ניתוח של נושא נתון המתייחס למורכבותו של הנושא, ולנקוט עמדה מנומקת (30 דקות)
- החיבור הוא אופציונלי ונדרש על ידי 40% מהמוסדות. 55% מהנבחנים בוחרים לכתוב חיבור.

■ דוגמה לנושא חיבור

אנשי חינוך מתלבטים בשאלה האם יש להאריך את משך הלימודים בתיכון לתקופה של חמש שנים, נוכח הדרישות הגוברות והולכות, מצד מעסיקים ומוסדות אקדמיים, שהתלמידים ישתתפו בפעילויות חוץ-לימודיות וקהילתיות בנוסף לדרישה לציונים גבוהים. ישנם אנשי חינוך התומכים בהארכת משך הלימודים לחמש שנים, מפני שהם חושבים שהתלמידים זקוקים לזמן רב יותר כדי להשיג את כל המצופה מהם. מחנכים אחרים אינם תומכים בהארכת משך הלימודים לחמש שנים, מפני שהם חושבים שהתלמידים יאבדו עניין בבית הספר, והוספת שנה תעודד נשירה מן הלימודים.

האם לדעתכם יש להאריך את משך הלימודים בתיכון לחמש שנים?

נקטו עמדה בנוגע לשאלה זו בחיבורכם. אתם יכולים לכתוב על כל אחת מנקודות המבט שהוצגו לעיל או להציג נקודות מבט שונה בשאלה זו. בססו את עמדתכם בעזרת נימוקים ודוגמאות ממוקדות.

GMAT (מנהל עסקים – תארים מתקדמים)

- בעבר כלל המבחן דרישה לכתיבת שני חיבורים, שלכל אחד מהם מוקצות 30 דקות:
1. ניתוח נושא – הבעת דעה מנומקת על נושא נתון מתחומי הלימוד של מנהל עסקים
 2. ניתוח טיעון – התייחסות אנליטית לטיעון נתון מתחומי הלימוד של מנהל עסקים
- במתכונתו החדשה של המבחן יש דרישה לניתוח טיעון בלבד.

■ דוגמה לניתוח נושא

"לעתים קרובות אנשים מתלוננים כי מוצרים אינם מיוצרים כך שיהיו עמידים. הם חשים כי ייצור מוצרים שמתבלים במהירות יחסית מבזבז משאבים אנושיים וטבעיים כאחד. עם זאת, הם אינם מבינים שפרקטיקות ייצור אלו מורידות את המחירים לצרכן ומעודדות את הדרישה למוצרים."

איזה טיעון משכנע יותר בעיניכם: התלונה על המוצרים הלא עמידים, או שמא התגובה לתלונה זו? הסבירו את עמדתכם אגב הצגת סיבות מתאימות ו/או דוגמאות מניסיונכם האישי, מתצפיות או מטקסטים שקראתם.

■ דוגמה לניתוח טיעון

הקטע הבא הופיע במדור המאמרים של ירחון חדשות ניהול:

"רוב החברות העסקיות יסכימו שככל שגדל הסיכון לפגיעה פיזית בעובד בעת מילוי תפקידו, כך צריך לגדול שכרו של העובד. מכאן שעבור המעסיקים יש היגיון כלכלי בהפיכת מקום העבודה לבטוח יותר: הדבר יפחית את הוצאות השכר שהם משלמים ויחסוך להם כסף."

דונו בהיגיונו של טיעון זה. הקפידו לנתח את קו המחשבה ולהשתמש בעדויות תומכות בטיעון. תוכלו לבדוק למשל אם ניתן להטיל ספק בהנחות היסוד שעומדות בבסיס הטיעון ולהציג הסברים חלופיים או דוגמאות נגדיות שעשויים להחליש את המסקנה. תוכלו גם להתייחס לסוג העדויות שיחזקו או יחלישו את הטיעון, לשינויים בטיעון שיוכלו להפוך אותו למבוסס יותר ולגורמים שיעזרו לכם להעריך את המסקנה שבסופו בצורה טובה יותר.

GRE (תארים מתקדמים)

המבחן כולל דרישה לכתיבת שני חיבורים:

1. ניתוח נושא – הבעת דעה מנומקת על נושא נתון (45 דקות)
2. ניתוח טיעון – התייחסות אנליטית לטיעון נתון (30 דקות)

■ דוגמה לניתוח נושא

ככל שבני האדם יסתמכו על הטכנולוגיה כדי לפתור בעיות, יכולתם לחשוב באופן עצמאי תלך ותידדרר.

דונו במידת הסכמתכם או אי-הסכמתכם עם הטיעון לעיל והסבירו את ההיגיון שבבסיס עמדתכם. בעת הצגת עמדתכם והטיעונים בעדה, עליכם לשקול את המובנים שבהם הטיעון עשוי להיות נכון או שגוי ולהסביר כיצד השיקולים הללו מעצבים את עמדתכם.

■ דוגמה לניתוח טיעון

בסקרים שנערכו בקרב תושבי מייסון סיטי התברר כי ענפי ספורט מים שונים (שחייה, שיט ודיג) הם בין פעילויות הנופש והפנאי החביבות עליהם. עם זאת, התברר כי לעתים נדירות בלבד משמש נהר מייסון, העובר בתוך העיר, לפעילויות אלו, וכי מחלקת הגנים העירונית מקדישה רק חלק קטן מהתקציב שלה לתחזוקת מתקני הנופש לאורך גדות הנהר.

במשך שנים התקבלו תלונות רבות מצד התושבים לגבי איכות מי הנהר והריח העולה ממנו. בתגובה לתלונות, הכריזה המדינה לאחרונה על תכנית לניקוי נהר מייסון. נראה אפוא שהשימוש בנהר לשם פעילויות ספורט מים יגדל. על כן על הרשויות העירוניות להקדיש כספים רבים יותר מתוך התקציב השנתי למתקני הנופש שעל גדות הנהר.

כתבו תגובה ובה תבחנו את ההנחות המוצהרות ו/או הסמויות שעומדות בבסיס הטיעון המובא לעיל. ודאו כי אתם מסבירים כיצד תלוי הטיעון בהנחות הללו והסבירו מה יהיו התוצאות אם יתגלה כי ההנחות הללו שגויות.

LSAT (משפטים)

הגריין במטלת הכתיבה מציג מצב ובו שתי אפשרויות בחירה. על הנבחן לבחור באחת מהן ולנמק את בחירתו. למבחן מוקצות 35 דקות.

(החיבור אינו מוערך, אלא נסרק ונשלח לאוניברסיטה לצד ציון המבחן הכללי. רוב האוניברסיטאות אינן משתמשות במבחן הכתיבה באופן סדיר.)

■ דוגמה

בני הזוג דוסון, שפרשו זה מכבר לגמלאות, שואפים לעבור מאזור עירוני גדול שבו חיו ועבדו במשך שנים רבות לקהילה קטנה יותר. הם עומדים למכור את ביתם, שבו שלושה חדרי שינה, ומתכננים לקנות דירה קטנה יותר, שבה שני חדרי שינה, במקום החדש שיעברו אליו. הם שוקלים שני יעדים אפשריים.

כתבו טיעון התומך באחת משתי האפשרויות. שים לב לשני קווים מנחים:

(1) הכנסתם החודשית של בני הזוג דוסון נמוכה כעת במידה ניכרת מהכנסתם בשנים שבהן עבדו. משום כך, ברצונם להקטין את הוצאות המחיה שלהם.

(2) בני הזוג דוסון מעוניינים לבלות את תקופת הפרישה שלהם בעיסוק במגוון רחב של פעילויות תרבות ופנאי.

העיר הייבן היל היא אתר נופש חורפי פופולרי, הממוקם במרחק שעתיים נסיעה מהאזור העירוני הגדול הקרוב ביותר. במשך חודשי החורף מגיעים תיירים מכל העולם להייבן היל כדי לגלוש במורדות ובמישורים המושלגים, ובעונה הבלתי מתוירת והשקטה יותר יכולים התושבים ליהנות ממסלולי הטיולים ומדיג באזור. הייבן היל מתגאה בסצנת האמנויות התוססת שלה ומארחת פסטיבל קולנוע שנתי קטן שהולך וגדל. המחיר הממוצע לדירה בהייבן היל דומה לערך הבית שמוכרים בני הזוג דוסון. במשך עונת החורף אפשר להשכיר לתיירים את הדירות בהייבן היל לתקופה קצרה במחיר גבוה במידה ניכרת מדמי השכירות בעונה הבלתי מתוירת.

בעיר ברוקוויל, הנמצאת במרחק של 45 דקות נסיעה מהאזור העירוני הגדול ביותר, נמצאת מכללת סוונסון – מוסד המתמחה במדעי הרוח והכולל חוג יוקרתי למוזיקה ולאמנויות הבימה. המכללה מארחת סדרת קונצרטים שנתית שמופיעים בה אמנים מקצועיים וידועי שם. כמו כן, המכללה מתגאה בספרייה ובגלריית אמנות מרשימות, ושתייהן פתוחות לציבור ללא תשלום. העיר עצמה היא עיר סטודנטית אופיינית, הכוללת תשתית רחבה של מסלולי אופניים ומגוון גדול של מסעדות וחנויות זולות, הפונות לתלמידי המכללה ולתושבים כאחד. מכללת סוונסון מציעה בימים אלו מגוון קורסי העשרה חוץ-אקדמיים זולים, הפתוחים לתושבי ברוקוויל. המחיר הממוצע של דירה בברוקוויל נמוך בהרבה מהבית שמוכרים בני הזוג דוסון.

MCAT (רפואה)

המבחן כולל שתי מטלות כתיבה. בכל אחת מוצג טיעון, והנבחן נדרש לכתוב חיבור המתייחס לטיעון (למשל הדגמה וניתוח). שתי המטלות שקולות באופיין ובזמן המוקצה להן: 30 דקות כל אחת.

■ דוגמאות

1

טבעו של המשטר הדמוקרטי הוא שהאזרחים בו תלויים זה בזה.

תארו סיטואציה ובה, מסיבה מוצדקת, אזרחים במדינה דמוקרטית לא יהיו תלויים זה בזה. דונו בגורמים הקובעים, לדעתכם, מתי אזרחים במשטר דמוקרטי צריכים להיות תלויים זה בזה.

מכל אמצעי התקשורת, לטלוויזיה כוח ההשפעה הרב ביותר על דעת הקהל.

תארו סיטואציה ובה הטלוויזיה איננה בעלת כוח ההשפעה הרב ביותר על דעת הקהל. דונו בגורמים הקובעים, לדעתכם, אם אכן לטלוויזיה כוח ההשפעה הרב ביותר על דעת הקהל.

בחינות בישראל

פרויקט מבחן הכתיבה של מאל"ן

פרויקט זה נערך בשנים 1995-1997, בעקבות החלטה של הנהלת המרכז הארצי לפתח מבחן כתיבה שיוצק לאוניברסיטאות כתוספת לבחינה הפסיכומטרית, ורוכז על ידי דנה גנאור-שטרן. בשל עלות גבוהה והיעדר ביקוש הוחלט לבסוף להקפיא את פיתוח המבחן. המבחן כולל שלוש מטלות כתיבה שלכל אחת מהן מוקצות 30 דקות, הנבדלות זו מזו בשני ממדים:

1. סוג הגריין שהמשימה מבוססת עליו – קטע קריאה אחד / שילוב שני קטעי קריאה / נתונים המוצגים בטבלה או בתרשים.
 2. סוג המטלה הנדרשת – סיכום / מתן הסבר / הבעת עמדה / יישום.
- בנוסף, יש שונות גדולה באורכי הגריינים (בין שליש עמוד לעמוד), אך לא באורך החיבור הנדרש – כעמוד אחד בכל המטלות.

■ דוגמה למטלה המבוססת על קטע קריאה

יחסי רופא-חולה – על פי "אתיקה ורפואה" מאת דוד הד

בסוגיה של יחסי רופא-חולה יש שתי גישות: הגישה הפטרנליסטית והגישה החוזית. המושג "פטרנליזם" נגזר מהמילה "אב". יחסי אב ובנים, או הורים וילדים, הם יחסים בלתי-סימטריים מובהקים. הגישה הפטרנליסטית מדגישה את חוסר הסימטריה ביחסים שבין רופא לחולה: הרופא הוא מספק השירות והחולה הוא הנוקד לו. היתרון העצום של הרופא על החולה ברמת הידע לא רק מזכה, אלא גם מחייב אותו לכפות את עמדותיו. תלותו של החולה ברופא היא מוחלטת, שכן היא נוגעת למצוקותיו הבסיסיות והדוחקות ביותר, ולכן יש צורך באמון בסיסי של החולה ברופא. הן יכולתו של החולה לבקר את התנהגות הרופא והן רצונו לעשות כן מוגבלים על ידי הכבוד שהוא רוחש למומחיותו ולמעמדו של הרופא.

הגישה החוזית יוצאת מן ההנחה שהטיפול הרפואי מושתת על יחסי הסכמה בין החולה לרופא. החולה הוא המבקש לקבל את שירותיו המקצועיים של הרופא, ויתרונו המקצועי של הרופא אינו מקנה לו את הזכות להתעלם מהעמדות הערכיות השונות והעצמאיות של החולה. אמנם בידי הרופא יתרון מכריע בכל הנוגע להכרת העובדות הרלוונטיות למחלה, אולם בידי החולה הזכות הבלעדית על גופו ועל חייו, ולכן גם על ההחלטה מה יעשה בפועל בגופו. מקור הסמכות של הרופא הוא רצונו של החולה, ולכן הסכמתו של החולה היא תנאי לכל טיפול בו.

בקטע מתוארות שתי גישות בנוגע ליחסי רופא-חולה. האם ניתן להצביע על שתי גישות דומות בנוגע ליחסי מורה-תלמיד? תארו את היחס שיעניק מורה הדוגל בגישה הפטרנליסטית ואת היחס שיעניק מורה הדוגל בגישה החוזית. מהי, לדעתך, הגישה העדיפה לתלמיד? מדוע?

היקף הכתיבה הנדרש: 20-30 שורות.

■ דוגמה למטלה מבוססת על שילוב שני קטעי קריאה

על מדע וערכים

משימת שילוב קטעים

קטע א: לפי "בין מדע לערכים ואידאולוגיה", מאת ישעיהו ליבוביץ'

הנחת היסוד של מאמר זה היא שקיימת אי-שייכות מוחלטת של המדע לערכים ולאידאולוגיה; אי-שייכות זו היא דו-סטרית. היא קיימת לא רק במובן זה שמדע אינו מבסס ערכים, אלא גם במובן זה שעולם הערכים והאידאולוגיה אין לו שום מגע עם המדע. התוצאות והמסקנות של העיסוק המדעי והמתודה המדעית אינן תלויות בטיב האדם המשתמש במתודה זו, כלומר באופי החינוך שקיבל, בערכים החברתיים שהוא מחזיק, ביחסיו החברתיים, באמונותיו, תקוותיו ושאיפותיו. המתודה המדעית היא אובייקטיבית במובן זה, שאין היא תלויה באדם המשתמש בה, אלא רק במידת כושרו להבין את המתודה הזאת לעצמה. בהינתן נתונים והמתודה המדעית כל החוקרים יגיעו לאותן המסקנות, ללא קשר לערכים המוסריים ולאידאולוגיה שבה הם מחזיקים.

קטע ב: לפי "המדע, המוסר והלימודים ההומאניים", מאת צ.א. קורצווייך

לפי ברונובסקי, הדבר המאפיין את דרכו של החוקר המדעי היא שאיפתו ומסירותו ללא סייג לאמת, ורק לאמת. החוקר הבודד אין ביכולתו להתחיל מחדש, ללא הנחות בדוקות ומבלי להסתמך על עבודתם של אלה שקדמו לו בחקר המדע. הוא סומך עליהם בידועו כי גם הם שאפו לחשוף את האמת. אם כי סומך המדען על אמיתות עבודות קודמיו, הרי לעתים הוא בודק אותן בדיקה ביקורתית, והוא אף עשוי להביע אי-הסכמה עם דעות מקובלות. היכולת לא להסכים ולחלוק על דעות מקובלות אפשרית רק באווירה של חופש. ולכן, חופש דיבור, חופש כתיבה וחופש הבעת דעה הכרחיים לאיש מדע כאוויר לנשימה. רואים אנו אפוא שמעצם העיסוק במדע אפשר לדלות מערכת שלמה של ערכים מוסריים, ואלה הם: אהבת האמת, חופש דיבור ומחשבה, צדק, יראת כבוד לזולת ואי-תלות רוחנית.

בעוד שישעיהו ליבוביץ' טוען במאמרו כי "קיימת אי-שייכות מוחלטת של המדע לערכים" טוען ברונובסקי ש"מעצם העיסוק במדע אפשר לדלות מערכת שלמה של ערכים מוסריים".

א. כיצד ניתן ליישב את הפער בין שני הציטוטים הנ"ל?

ב. בחר תחום עיסוק נוסף (למשל: הוראה, רפואה, עיתונאות) והבע דעתך לגביו: האם קיימת אי-שייכות מוחלטת של התחום שבחרת לערכים? האם מעצם העיסוק בתחום שבחרת אפשר לדלות מערכת שלמה של ערכים מוסריים?

היקף הכתיבה הנדרש: כ-30 שורות.

■ דוגמה למטלה מבוססת טבלה

במחקר שבדק את מקורות התמיכה שבהם נעזרים בני נוער הוצגה בפני תלמידי כיתות ט' ו-י"א השאלה הבאה:

"אם יש לך בעיה המטרידה אותך, למי תפנה קודם כל כדי לקבל עצה?"

בטבלה שלפניכם מוצג אחוז הבודדים בכל מקור עצה מתוך הקבוצה הנשאלת. הנתונים מוצגים בנפרד לכל שכבה (כיתות ט' מול כיתות י"א), ובתוך כל שכבה הנתונים מוצגים לחוד עבור בנים ועבור בנות.

טבלה 1: מקורות העצה של בנים ובנות מכיתות ט' ו-י"א (באחוזים)

מקור העצה	כיתות ט'		כיתות י"א	
	בנים	בנות	בנים	בנות
חברים	50%	70%	73%	80%
אח/אחות	11%	8%	6%	11%
הורים	25%	17%	10%	6%
מורים, יועצים	1%	0%	1%	0%
לא פונים לבקש עצה	13%	5%	10%	3%
סך הכול	100%	100%	100%	100%

משימות הכתיבה

א. מהן המגמות העיקריות המסתמנות מנתוני הטבלה? התייחסו להבדלים בין המינים ולהבדלים בין השכבות.

ב. בחרו ממצא אחד המעניין אתכם במיוחד והסבירו על מה לדעתכם הוא מעיד.

היקף הכתיבה הנדרש: 25 שורות.

[הערה: משימות אלה הועברו העברה ניסיונית על תלמידי המכינה הקדם-אקדמית באוניברסיטה העברית בירושלים, נבדקו בהערכה הוליסטית ואנליטית, ועמדו בקריטריונים לשימוש בבחינה].

בחינת הבגרות בחיבור

בבחינה נדרש לכתוב חיבור, ועל הנבחן לבחור נושא אחד מתוך שישה נושאים. בארבעה-חמישה מהנושאים הגריין כולל קטע קריאה באורך חצי עמוד עד עמוד, ובנושא אחד או שניים הגריין קצר מאוד: ניתנת בו מובאה ופירושה, או אמירה קצרה, והנבחן מתבקש להתייחס אליהן. היקף הכתיבה הנדרש הוא 2-3 עמודים ולכתיבת החיבור מוקצות 3 שעות.

■ דוגמה – גריין ארוך

הזמר העברי כראי לערכי החברה

החגיגות לכבוד 60 שנים למדינת ישראל שבו והעלו לתודעת הציבור את שירי ארץ-ישראל, המלווים את המדינה בכל תקופותיה. בקטע שלפניך טוענת הכותבת שהשינויים בזמר העברי מבטאים שינויים בערכי החברה בישראל.

כתוב מאמר והגב בו על טענה זו. הסבר כיצד הזמר העברי משקף את ערכי החברה בישראל בזמננו, והדגם באמצעות שירים מוכרים לך.

(מעובד על פי מאמר מאת ד"ר טלילה אלירם, מפנה, ספטמבר 2007)

שירי ארץ-ישראל עוסקים הן בחוויות אישיות והן בחוויות משותפות לכלל החברה. אחד הנושאים המרכזיים בשירי ארץ-ישראל הוא היחס למולדת, לארץ-ישראל. בתקופת העליות הראשונות הפליגו השירים בשבח הארץ, שפעו אופטימיות והביעו אהבת מולדת ונכונות להקרבה. הם היו ספוגים בלהט העשייה ובחדוות היצירה. לנחישות, לביטחון, לאומץ ולשמחה שבמילות השירים היה ביטוי גם בלחנים, שהם בעלי מקצב של מרש או של ריקודי הורה.

מאין נבעו הנחישות והמסירות העולות מן השירים של תקופה זו? לדעת הוגה הדעות נתן רוטנשטרייך, התאפיין היישוב בארץ-ישראל בראשיתו במחויבות עמוקה למטרה משותפת, שהיא תנאי להשגת החזון הציוני: שאיפה לתחייה לאומית עברית בארץ-ישראל ובניית חברה הנשענת על ערכי הסוציאליזם. מחויבות זו, שהתקיימה גם בשנותיה הראשונות של המדינה, סייעה בגיבוש חברה בעלת חזון נפשי ועוצמה רוחנית.

התרבות המוזיקלית שנוצרה באותן שנים, שירי שיבולים ורועים ושירי נופים, ביטאה גם אידאולוגיה של קרבה לאדמה ורצון להידמות לעברי הקדמון.

במהלך העשור השני והשלישי למדינה חלו בחברה הישראלית שינויים אידאולוגיים ותרבותיים. הערכים החברתיים כמו סוציאליזם, שיתופיות וסולידריות החלו לפנות את מקומם לערכים כמו תחרותיות, הישגיות ונהנתנות. שינויים אלה מצאו את ביטויים גם בתרבות המוזיקלית בישראל והשפיעו על שירי ארץ-ישראל, כפי שאמר אהוד מנור (בריאיון בשנת 1996):

"השיר האישי נעשה רווח בעולם כולו. התופעה הבולטת ביותר בזמר היא שחדלנו לשיר בגוף ראשון רבים. לא עוד 'אנחנו שרים לך מולדת ואמא', אלא 'נתתי לך חיי, ירדתי על ברכיי'. היחס השתנה, כי החלו להיבקע סדקים בחברה הישראלית".

■ דוגמה – גריין קצר

רשלנות פושעת

"גם מתרפה בקלאקתו, אֵח הוא לבעל משחית". (משלי יח ט)

אחד הפירושים לפסוק זה הוא שאדם המתרשל במלאכתו, מזלזל במשימה המוטלת עליו ומוסר העבודה שלו נמוך – דומה לאדם שמחבל ומשחית בזדון.

כתוב מאמר והסבר בו את הטענה העולה מפירוש זה, הבע את דעתך עליה, ונמק אותה בעזרת דוגמאות מתחומים שונים (לדוגמה: רפואה, תחבורה, בנייה, ביטחון).

חלק ב – פיתוח מטלת הכתיבה במכפ"ל

תהליך פיתוח מטלת הכתיבה – החלטות ראשוניות

בעקבות בחינת הסוגיות המרכזיות בתהליכי פיתוח מטלות כתיבה וסקירת המצב בבחינות מקבילות למכפ"ל, ונוכח האילוצים המעשיים, התקבלו כמה החלטות הנוגעות לשלבי הפיתוח הראשוניים של המטלה:

ניתוח תוכן

הוחל בניתוח יכולת הכתיבה ובמיפוי של מרכיביה, כדי לסייע בבניית מטלה שייוצגו בה כל המרכיבים הרלוונטיים, ובפיתוח מחוון שישקף אותם. רשימת המרכיבים נאספה ממחקרים בנושא, וכבסיס ראשוני לה שימש מערך התכונות המופיע במחוון בחינת הכתיבה שפותחה בעבר במאלי"ו. התהליך מוצג ביתר פירוט בחלק ג להלן.

בחינה ישירה של יכולת הכתיבה

בשל מגבלת הזמן הקיימת במכפ"ל (כ-30 דקות לחלק הבוחן כתיבה) אי אפשר להשתמש בבחינה במטלות משני הסוגים – ישירה ועקיפה – ויש לבחור באחד מהם. אף שלמטלות כתיבה עקיפות כמה יתרונות ניכרים, ובראשם מהימנות גבוהה מזו של מטלות ישירות, אף בחינה מבחינות המיון שנסקרו אינה נסמכת על פריטים עקיפים בלבד, ובבחינות הכוללות פריטים אלה (SAT, ACT) מתלווה אליהם מטלת כתיבה ישירה. נראה כי האותנטיות הנמוכה יחסית של פריטים עקיפים אינה מאפשרת להסתפק בהם במבחני כתיבה. בנוסף, נראה כי תוקפה ההשלכתי של בחינה עקיפה נמוך מזה של הבחינה הישירה: כמה מחקרים מצאו קשר בין הנהגת מבחני כתיבה עקיפים ובין צמצום היקף לימודי הכתיבה, והחלפתם בהכנה ממוקדת למטלות הכתיבה העקיפות (Murphy, 2003), ולעומת זאת, הכללת מטלת כתיבה ישירה ב-SAT הביאה לתגבור הוראת הכתיבה בבתי הספר ולהגדלת מספר מטלות הכתיבה המוטלות על התלמידים. שיקולים אלה מכתיבים בחירה במבחן כתיבה ישירה, על אף מגבלותיו.

התמקדות בכתיבה עיונית-טיעונית

על מטלה המשמשת להערכת יכולת כתיבה ומיועדת לסייע במיון מועמדים למוסדות אקדמיים להיות אותנטית ככל האפשר, כלומר דומה באופייה דמיון מרבי למטלות הכתיבה באוניברסיטה. סוגת הכתיבה הרווחת במוסדות האקדמיים היא כתיבה עיונית – כתיבה המשמשת להצגת רעיונות ולדיון בהם. "רעיון" הוא דעה, תפיסה, גישה או תופעה כלשהי, והדיון בו יכול לכלול פירוט, הסבר, נימוקים משכנעים, הוכחות, מסקנות וכולי. מבנה הטקסט העיוני אמור לשקף את אופן הפיתוח של הרעיון: על חלקיו השונים להיות מקושרים זה לזה בדרך הגיונית, ועל היחס בין הטיעונים השונים להיות ברור. כתיבה עיונית טובה מתאפיינת בלשון מדויקת ובהירה ובאחידות משלבית. כתיבה עיונית, במיוחד כזו המשמשת במוסדות אקדמיים, מבוססת על מקורות ותלויה בהבנתם.

כאמור, מבין דגמי מטלות הכתיבה הקיימים, לשלושה רמת אותנטיות גבוהה בכל הנוגע להערכת יכולת כתיבה אקדמית-עיונית: התייחסות ביקורתית-אנליטית לקטע, בניית מאמר מתוך מקורות ונקיטת עמדה מנומקת בנוגע לקטע. מבין אלה, שני הראשונים דורשים הצגת גריינים או מקורות מפורטים וארוכים מכדי להתאים למסגרת הזמן המוקצב למטלה במכפ"ל. הדגם השלישי – נקיטת עמדה מנומקת הנוגעת לגריין נתון – אפשרי גם עם גריינים קצרים יחסית, דבר המאפשר שמירה על רמת אותנטיות גבוהה בתוך מסגרת הזמן הקיימת במכפ"ל. לפיכך הוחלט שהמטלות שיוצגו בשלב הפיתוח הראשוני יהיו מטלות כתיבה עיונית-טיעונית – מטלות הדורשות נקיטת עמדה בסוגיה המוצגת בהן. בנוסף לדרישה לכתיבה עיונית הנגזרת מהמטלות עצמן, הוחלט להנחות את הנבחנים במפורש לכתיבה מסוג זה, ולהציג להם במפורט את מאפייניה.

מידת ההתבססות על הבנת הנקרא

כאמור, מרבית הגישות התיאורטיות והממצאים המחקריים מראים קשר הדוק בין יכולת הכתיבה ובין יכולת הקריאה. רבים טוענים כי מסיבה זו אין ללמד קריאה וכתיבה בנפרד ואין לבחון אותן בנפרד. מגישה זו עולה, שהמשימות במבחנים המיועדים לבחון יכולת כתיבה צריכות להיות מבוססות על קטעי קריאה. משימות כאלה יהיו גם אותנטיות יותר, שכן כאלה הן רוב רובן של המטלות באוניברסיטה. מנגד, השאיפה לבדוד ככל האפשר את יכולת הכתיבה מיכולות אחרות שנמדדות באופן בלתי תלוי בבחינה – ובפרט מיכולת הבנת הנקרא, שמוקדש לה חלק נכבד מתחום החשיבה המילולית במכפ"ל – מצדיקה את צמצום התלות במרכיב הקריאה במטלה, גם במחיר פגיעה באותנטיות. הוחלט לבחון את שתי האפשרויות, ולכלול בבדיקה הן מטלות שבהן התלות ביכולת הבנת הנקרא גבוהה והן מטלות שבהן היא נמוכה.

מידת ההתבססות על ידע קודם

מכיוון שמיון המועמדים לאוניברסיטאות נעשה כדי לבדוק את יכולתם של המועמדים ולא את הידע הקודם שלהם, ובהינתן שעל המטלות להיות אותנטיות ככל האפשר אגב שמירה על מסגרת זמן מוגבלת, הוחלט שהמטלות המתאימות ביותר לפיתוח ראשוני הן מטלות משני הסוגים האלה: מטלות המבוססות על גריין כללי ומינימלי, שמטבעו דורש ידע כלשהו לביצוע המטלה, ומטלות המתבססות על מידע נתון (קטע קריאה), המצמצם את ההשפעה שיש לידע קודם בעולם התוכן על ביצוע המטלה.

תפקיד הגריין במטלה

בשלב הראשון של הפיתוח הוחלט לבחון הן מטלות המתייחסות ישירות לגריין והן מטלות שהגריין משמש בהן רק הקשר.

מידת המיקוד של המטלה

בשלב הראשון של הפיתוח הוחלט לבחון רמות שונות של מיקוד: הן מטלות עם נושא פתוח יחסית והן מטלות עם נושא ממוקד.

היבטים מעשיים בפיתוח מטלת הכתיבה

לשאלות העקרוניות הנוגעות לפיתוח מטלת הכתיבה נוספו כמה מאפיינים מעשיים של המטלה, שנגזרים מהתנאים שהיא מועברת בהם (מאפייני מכפ"ל והטכנולוגיה העומדת לרשותנו בשלב זה):

מסגרת זמן של כ-30 דקות

כדי לשמור על מסגרת הזמן של מכפ"ל קוצרו פרקי הבררה שבו ב-40 דקות בסך הכול (20%). פירוש הדבר הוא מגבלת זמן של כ-30 דקות למטלת הכתיבה (היתרה מוקדשת להקראת הוראות נוספות, איסוף החיבורים ומעבר להמשך הבחינה).

החיבורים נכתבים על נייר

לעת עתה מכפ"ל הוא מבחן נייר ועיפרון (למעט מקצת הנוסחים המיועדים לנבחנים בתנאים מותאמים ומועברים במחשב). עם פיתוח תשתית טכנולוגית מתאימה יישקל מעבר לכתיבת חיבורים במחשב לכלל נבחני מכפ"ל.

למגבלות תלויות הנסיבות שיכולת הכתיבה במבחן מוערכת בהן, ובראשן מגבלת הזמן, יש גם משמעות תאורטית: הן פוגעות בכל אחד משלושת שלבי הכתיבה במודל הכתיבה שהציגו Flower & Hayes (1981). שלב הביקורת נפגע יותר מכול, אולם גם שלב התכנון נפגע במידת מה, ובמידה פחותה גם שלב יצירת הטקסט. לכן הוחלט להתייחס לתוצרי הכתיבה במבחן כאל טיוטה ראשונה, ולא כאל תוצר מוגמר, הממצה את יכולת הכתיבה של הכותב. ברלנד ועמיתיו (Breland et. al., 1999) מציגים ממצאי מחקר שמעידים על קיומו של קשר בין הזמן המוקצב לכתיבת חיבורים ובין ממוצע הציונים בהם: הוספת זמן לכתיבת החיבורים הביאה לעלייה מתונה בציוני החיבורים. בבחינות ה-TOEFL למשל נמצא הבדל ממוצע של 0.3 נקודות בסולם הוליסטי של 0-6 נקודות, בחיבורים שהוקצו להם 45 דקות, לעומת חיבורים שהוקצו להם 30 דקות. עלייה דומה נרשמה בחיבורים במבחני ה-SAT II וה-Subject GRE. עם זאת, הם מציגים גם עדויות מצטברות המצביעות על כך שמדרג הציונים אינו משתנה עם הארכת הזמן. בעת פיתוח מטלת הכתיבה ופורמט הבחינה החדש נבדקו גם במאל"ו ההבדלים בין הערכות חיבורים בהקצאות זמן שונות (30 דקות, 35 דקות, 40 דקות ו-45 דקות). נמצא שהן מדרג הציונים הן ממוצעייהם לא הושפעו מתוספת הזמן.

פיתוח דגמי המטלות הבסיסיים

נוכח ההחלטות שנבעו מבחינת הסוגיות העקרוניות, מסגרת האילוצים המעשיים וסקירת המבחנים המקבילים למכפ"ל, פותחו שני סוגי מטלות, ומכל אחד מהם נגזרו שני דגמים.

מטלות כתיבה המבוססות על גריינים ארוכים

על מבחן הערכת יכולת כתיבה שמטרתו לסייע במיון מועמדים למוסדות אקדמיים להתמקד בהערכת מיומנויות הכתיבה הנדרשות במוסדות אלה. עיקר הכתיבה במוסדות האקדמיים היא עיונית: היא מבוססת על מקורות ותלויה בהבנתם. לפיכך, ביסוס מבחן הכתיבה על טיפול

בהבנת מקורות תורמת לאוטנטיות שלו. נוסף על כך, היא מקטינה את משקל הידע הקודם הנדרש לביצוע הבחינה.

דגם 1 – מטלה המבוססת על גריין ארוך אחד

הדגם הראשון מסוג זה מבוסס על גריין ארוך אחד, ואורך החיבור הנדרש בו הוא 20-30 שורות.

■ דוגמה

על שפה וחברה – מתוך מאמר מאת שולמית הראבן

ברצוני להציג שאלת תם: מה פירוש הביטוי הרווח "נשים מוכות"? זו מין תכונה כזאת, מולדת או נרכשת? זה גורל? זו מחלה? כשם שיש נשים חולות סוכרת, לא עלינו, כך יש גם נשים מוכות? האם מדובר בסוג של מצב נשי-ספציפי שאין איש אחראי לו? הבה נאמר זאת בבירור: אין דבר כזה, בעיה של נשים מוכות. זאת לא הבעיה שלהן. יש רק גברים מכים. עצם השימוש במונח הרווח מסיר באחת את האחריות מן הגבר הסדיסט האלים, וחוזר ומטיל אותה על האישה, כמעין סטיגמה: מה לעשות, היא שייכת לזן הנשים המוכות. ועד כדי כך הגיע האבסורד ששמענו על ארגונים ואנשי מקצוע המדברים על... "בעליהן של נשים מוכות". מסכנים באמת, התחתנו, לרוע המזל, עם נשים כאלו.

תהום רובצת בין השימוש במונח "נשים מוכות" לבין השימוש במונח "גברים מכים". בתפיסה שאותה כופה עלינו המונח "נשים מוכות", האישה היא שצריכה לעזוב את הבית, מפני שהיא הבעיה, ולא הוא. בתפיסה הזאת צריך להקים מקלט לנשים מוכות, כלומר להרחיק אותן מביתן ומילדיהן, כי הגברים נשארים בבית. הם חומקים איך-שהוא מן השפה, מן המינוח, מן האחריות – ואחר כך גם מן העונש הראוי, כי יש להם, במקרה ולרוע המזל, אישה מוכה. לו היינו משתמשים במונח המתאים – "גברים מכים", היה ברור שאין צורך במקלטים לנשים מוכות, כי אם בבתי מאסר לגברים מכים, ומהר.

מה לדעתכם טענתה המרכזית של שולמית הראבן? מה דעתכם על טענתה?

כמה מהמטלות מדגם זה מבוססות על מקורות – בדרך כלל קטעים מתוך מאמרים – שכמעט שלא שונו, כדי לשמור על אותנטיות גבוהה. תוצאה ישירה של הדבר היא שהיקף המידע החדש שמוצג לנבחנים ושניתן לשאול עליו מצומצם יחסית, ואינו מובא בהכרח בצורה הבהירה ביותר. במטלות אחרות המקורות מעובדים. מטלות אלה אותנטיות פחות, אבל מאפשרות הצגה בהירה יותר של מידע רב יותר.

דגם 2 – מטלה המבוססת על שילוב של שני גריינים

דגם זה מבוסס על גריין המורכב משני טקסטים נפרדים, ואורך החיבור הנדרש בו הוא 20-30 שורות.

קטע א: אישור תרופות חדשות

הרשות האמריקנית לתרופות ולמזון היא הגוף הממשלתי הבודק תרופות חדשות בסיום תהליך הפיתוח שלהן ומחליט אם לאשר את הפצתן בארצות הברית. הרשות דורשת מחברות התרופות המבקשות לקבל אישור הפצה לתרופה חדשה לערוך מחקרים רבים העומדים בקריטריונים מדעיים מחמירים. על הרשות לפעול באחריות, במקצועיות ובאמינות גבוהות ביותר, שכן החלטותיה יכולות להשפיע על חיים או מוות של בני אדם. לדוגמה, היא עשויה להידרש לשאלה אם לאשר הפצת תרופה חדשה לטיפול במחלה קטלנית אף שתופעות הלוואי של התרופה אינן ידועות מספיק.

יש הטוענים כי הפקדת הסמכות הבלעדית להחליט על שימוש בתרופות חדשות בידי רשות ממשלתית אינה מוצדקת, וכי במקרים מסוימים יש להעביר סמכות זו לחולים הזקוקים לתרופה – לשתף אותם במידע הקיים על התרופה, על יכולת הריפוי שלה ועל תופעות הלוואי שהיא גורמת, ולהשאיר בידיהם את ההחלטה אם להשתמש בה.

קטע ב: הגבלת סרטים אלימים

סרטים ובהם תיאור בוטה של אלימות מביאים, לדעת רבים, לידי התנהגות אלימה בקרב הצופים בהם. הטוענים זאת מסתמכים על מחקרים שהראו כי אצל בני אדם שמלכתחילה דיווחו על נטיות לאלימות, צפייה בסרטים שבהם תכנים אלימים גרמה לשחרור העכבות והגבירה את הסיכוי שיעשו מעשה אלים. לכן, לדעת רבים, יש הצדקה להטיל הגבלות מטעם השלטון על יצירה והפצה של סרטים מסוג זה.

מן הצד השני עולה השאלה אם אין הגבלה כזאת בידי השלטון מתנגשת עם העיקרון הדמוקרטי של חופש הביטוי. יש הסוברים כי על השלטון להימנע מלהגביל יצירה והפצה של סרטים המציגים תכנים אלימים, וכי יש לאפשר לכל פרט לבחור את סוג התכנים שהוא מעוניין להיחשף אליהם.

התערבות השלטון מול חופש הפרט: דונו במתח הקיים בין שני גורמים אלה, וכתבו – בנוגע לדוגמאות שהובאו בקטעים שלעיל – אם אתם מצדדים בהתערבות השלטון או מתנגדים לה, ומאילו שיקולים.

כאמור, יתרונותיהן של מטלות כתיבה המבוססות על גריין ארוך הם אותנטיות גבוהה וצמצום ההסתמכות על ידע קודם. חסרונותיהן הם זמן ארוך יותר הדרוש לקריאתן ותלות מובנית ביכולת הבנת הנקרא. בנוסף, מטלות שילוב הקטעים דורשות זמן עיבוד רב, בנוסף לזמן הקריאה.

מטלות כתיבה המבוססות על גריינים קצרים

מטלות המבוססות על גריינים קצרים נמצאות בשימוש במרבית מבחני המיון המקבילים למכפ"ל (SAT, ACT, GRE, GMAT, MCAT). במטלות אלה הזמן המוקדש לקריאת המטלה קצר ביותר, וגם התלות ביכולת הבנת הנקרא מינימלית. זאת במחיר ויתור על אותנטיות ועל הסתמכות על ידע קודם, וכן היעדר הקשר, שעלול להוביל לשונות גדולה מדי בפרשנות הגריין.

דגם 3 – מטלה המבוססת על גריין קצר אחד

אורכי הגריינים והמטלות בדגם זה נעו בין 2 ל-6 שורות, ואורך החיבור הנדרש הוא 20-30 שורות.

■ דוגמה

חבר כנסת גיבש הצעת חוק המטיל על הורים את האחריות לעברות שביצעו ילדיהם אשר טרם מלאו להם 18 שנים. לפי הצעת החוק הילדים יישפטו, ואם יימצאו אשמים יוטל על הוריהם קנס כספי.

מה דעתכם על החוק המוצע? נמקו את עמדתכם.

דגם 4 – שילוב של שתי מטלות המבוססות על גריין קצר

במהלך הפיתוח הועלתה גם האפשרות להציג במסגרת הזמן המוקצה למטלת הכתיבה (30 דקות) שתי מטלות קצרות, שלכל אחת מהן יוקצו 15 דקות. אורך החיבורים שיידרשו בכל אחת מהמטלות יהיה 10-15 שורות. מטרת הפיתוח של דגם זה הייתה לבדוק האם הכפלת מספר הפריטים יכולה לתרום למהימנות של מדידת יכולת הכתיבה. עם זאת, השימוש בדגם זה מעורר את החשש שמגבלת הזמן דווקא תשטיח הבדלים בין נבחנים (מכיוון שהחיבורים שייכתבו יהיו קצרים מכדי לספק אומדן טוב ליכולת הכתיבה), ושלאחר הזמן יהיה גדול מדי וייתפס כמוגזם וכלא הוגן. גם בדגם זה נעו אורכי הגריינים והמטלות בין 2 ל-6 שורות.

■ דוגמה

מטלה א:

חבר כנסת גיבש הצעת חוק המטיל על הורים את האחריות לעברות שביצעו ילדיהם אשר טרם מלאו להם 18 שנים. לפי הצעת החוק הילדים יישפטו, ואם יימצאו אשמים יוטל על הוריהם קנס כספי.

מה דעתכם על החוק המוצע? נמקו את עמדתכם.

מטלה ב:

הרצון החופשי של בני האדם הוא אשליה – גורלנו נקבע על פי הנסיבות שאליהן נולדנו.

מה דעתכם על טענה זו? נמקו את עמדתכם.

סדרת שאלות המתייחסות לגריין אחד

דגם נוסף שנבחן – ונפסל בשלב מוקדם יחסית של הפיתוח – התבסס על סדרת שאלות המתייחסות לאותו גריין. השאלות מוצגות בנפרד זו מזו, ולכל אחת ניקוד נפרד.

■ דוגמה

מחקרים שנעשו מטעם האו"ם מגלים תמונה עגומה של אוכלוסיות הדגים החיים במים בחופי צפון מערב אפריקה. מתברר שבאזור זה אוכלוסיית הדגים החיים במים עמוקים היא רבע ממה שהייתה לפני 25 שנה. מיני דגים מסוימים אחדים – המבוקשים מאוד במסעדות יוקרה או שדגים אותם למטרות מסחריות אחרות – נעלמו כליל או נמצאים בסכנת הכחדה.

הסיבה לתופעות אלה ברורה: דיג יתר. צי עצום של ספינות דיג ומכמורות תעשייתיות מאירופה, מסין וממדינות אחרות "מרוקן" את הים בצפון מערב אפריקה ומנשל את הדגים המקומיים מפרנסתם. דייגים אפריקנים מאשימים את האירופים בזלזול ובצריכת יתר, שיש בהן ביטוי לחוסר אכפתיות כלפי הדייגים האפריקנים הוותיקים.

פקידים מהאיחוד האירופי טוענים כי מדינות אפריקה נכשלו בביקוח על הדיג, ולכן הן אשמות במצב. לדבריהם, ממשלות אפריקניות תאבות בצע מוכרות לזרים זכויות דיג רבות יותר משניתן לספק ומאפשרות לספינות בלתי-מורשות לדוג באזורי רבייה.

1. הציגו במילים שלכם את הבעיות המרכזיות שמציג הקטע. (5 שורות)

2. יש הטוענים כי האירופים, בהתנהגותם חסרת האחריות, יהפכו בעתיד לקורבנותיה של תאוותנותם- שלהם. מה דעתכם על אמירה זו? נמקו את עמדתכם. (20 שורות)

לדגם כזה יש יתרונות: הוא כולל כמה פריטים (גם אם תלויים במידת-מה זה בזה), ומאפשר להציג גריין עשיר וארוך יחסית. אף על פי כן, הוא נפסל בשלב מוקדם, בעיקר בשל הדמיון הרב שלו לפריטי הבנת הנקרא שנכללים כבר במכפ"ל, אך גם בגלל חסרונות נוספים: היעדר הקצאת זמן מדויקת לכל פריט, קבלת תשובות מוגבלות באורכן, שאינן בהכרח מייצגות לעומק את כל מרכיבי תהליך הכתיבה, וקושי ביצירת נוסחים שקולים, בגלל חלוקת שאלות שונה, שנגזרת מאופיו של כל קטע וקטע.

ניסוי ראשון

לצורך הערכת הדגמים השונים נבדקו 12 מטלות – שלוש מטלות מכל אחד מארבעת הדגמים. בשלב מקדים נבדקו המטלות על כעשרים מעובדי מאל"ו. בדיקות אלה שימשו לפיתוח המטלות ולשיפורן, וכן לגיבוש ראשוני של ההוראות לבוחנים, של ההנחיות לנבחנים, של היקף הכתיבה הנדרש ושל עיצוב המטלה ודף הכתיבה.

12 המטלות הועברו כחלק מניסוי שבדק את הפורמט החדש של הבחינה כולה. בניסוי השתתפו 742 נבדקים בשפה העברית (62 בממוצע למטלה) ו-628 נבדקים בשפה הערבית (52 בממוצע למטלה). המשתתפים בניסוי היו נבחנים שעמדו להיבחן במכפ"ל כשבוע לאחר העברת הניסוי.

המטלות הוצגו למשתתפים בניסוי בתחילת הבחינה, כפרק נפרד, והחיבורים נכתבו על דפי תשובות נפרדים. הנבחנים הונחו להשתמש בדפי החוברת אם יצטרכו דפי טיוטה. בסוף הבחינה מילאו הנבחנים דפי משוב ובהם נשאלו שאלות הנוגעות למטלת הכתיבה (היקפה, מגבלת הזמן, התאמתה לבחינה וכולי). כדי לשמור על רמת מוטיבציה גבוהה בעת ביצוע מטלת הכתיבה הוגרלו פרסים כספיים בין הנבחנים שחיבוריהם היו בעלי ציון ממוצע ומעלה.

תוצאות הניסוי הראשון

טבלה מספר 1 – מתאמים בין הערכות שני המעריכים, על פי דגם המטלה, $n=742$

דגם	מתאם בין הערכות המעריכים
גריין קצר בודד ($n=212$)	0.48
צמד גריינים קצרים ($n=197$)	0.51
גריין ארוך ($n=175$)	0.61
שילוב קטעים ($n=158$)	0.53

יש לציין כי היכולת להכליל מהממצאים מוגבלת, בין השאר משום שכל מטלה הוערכה על ידי זוג מעריכים אחר (כלומר, ייתכן שלזהות המעריכים הייתה השפעה גדולה על המתאמים).

טבלה מספר 2 – השוואה בין ארבעת דגמי מטלות הכתיבה שנבדקו, על פי המתאמים בין ציוני מטלת הכתיבה ובין ציוני האמת בבחינה בתחום המילולי (VE) ובבחינה כולה (TOT). (n=718)

בכל סוג מטלה נבדקו מתאמים עבור ציון התוכן, עבור ציון הלשון ועבור הציון הכללי לחיבור (לפיתוח אופן הערכת החיבורים, ובכלל זה החלוקה להערכת תוכן והערכת לשון, ראו חלק ג להלן).

		תוכן	לשון	כללי
גריין קצר בודד (n=206)	VE	0.32	0.42	0.39
	TOT	0.29	0.43	0.38
צמדי גריינים קצרים (n=190)	VE	0.26	0.45	0.38
	TOT	0.25	0.41	0.36
גריין ארוך (n=170)	VE	0.39	0.41	0.44
	TOT	0.39	0.43	0.44
שילוב קטעים (n=152)	VE	0.14	0.20	0.18
	TOT	0.22	0.27	0.26

יש פער יציב לטובת המתאמים עם ציון הלשון, לעומת המתאמים עם ציון התוכן, שנע בין 0.19 ל-0.03 במתאמים עם התחום המילולי, ו-0.16-0.04 עם הציון בבחינה כולה (גם בניסוי השני, שיוצג בהמשך, נמצא פער באותו כיוון, אם כי קטן יותר). הפער הגדול ביותר בין המתאמים עם הציון בלשון ובין המתאמים עם הציון בתוכן היה במטלות צמדי הגריינים הקצרים – כנראה מפני שקשה לפתח תוכן של ממש במסגרת של 15 דקות.

יש כמה סיבות אפשריות לפערים בין מתאמי הציונים בבחינה עם ציון התוכן ובין המתאמים עם ציון הלשון: ראשית, הערכת ממד הלשון נסמכת על כללים פורמליים וברורים יותר. הדומיננטיות של כללים אלה הייתה גדולה במיוחד בניסוי הראשון, שהבודקים השתמשו בו

במחווה ראשוני, כך שהמעריכים, שברובם היו בעלי רקע בהערכת לשון, נסמכו בהערכותיהם בעיקר על רקע זה. בנוסף, נראה כי תחום התוכן רחוק מהתכנים הנמדדים בפרקי הבררה יותר מתחום הלשון, ולכן הצלחה בו מתואמת פחות עם הצלחה בפרקים אלה. ממצא אחר שעלה בניסוי הוא שמתאמי מטלות שילוב הקטעים נמוכים במידה ניכרת ממתאמי המטלות האחרות. כאמור, יכולת ההכללה מהבדיקה מוגבלת, מכיוון שלכל אחת מ-12 המטלות הוצמד זוג מעריכים מיוחד לה.

מסקנות מהניסוי הראשון והיערכות לקראת הניסוי השני

בהסתמך על שילוב הנתונים הסטטיסטיים שהתקבלו, על משובי הנבחנים ויותר מכול – על עמדתם המקצועית של חברי צוות הפיתוח, התקבלו כמה החלטות הנוגעות להמשך תהליך הפיתוח: הוחלט לוותר על מטלות צמדי הגריינים הקצרים ועל מטלות שילוב הקטעים. עוד הוחלט לוותר על מטלות הגריינים הארוכים ועל מרבית הגריינים הקצרים, ולפתח במקומם מטלות שישלבו את היתרונות שבשני הדגמים, וזאת באמצעות עיבוי הגריינים הקצרים (עד 10 שורות, כולל המטלה) כדי להעלות את רמת האותנטיות שלהם, אגב הקפדה על לשון פשוטה ובהירה ככל האפשר, כדי להבטיח שהשפעתה של יכולת הבנת הנקרא על ביצוע המטלה תהיה מינימלית. כמו כן הוחלט לחזק את הזיקה בין המטלה ובין הגריין, באמצעות ויתור על מטלות שהגריין מספק בהן רק קונטקסט. לבסוף נוסחו עקרונות בסיסיים בחיבור מטלות.

ויתור על מטלות צמדי הגריינים הקצרים

לכאורה, מבחינת מהימנות הבחינה, שימוש בשתי מטלות קצרות עדיף משימוש במטלה אחת ארוכה. אולם נראה כי החיבורים הנכתבים בתגובה למטלות הקצרות הם מדגם כתיבה קצר מדי, והזמן הקצר המוקצב לכתיבתם עלול להקשות על הנבחן לבטא את יכולת הכתיבה שלו. בחיבורים רבים הפתיחה והסיכום היו שבלוניים – ובחיבור של 10-15 שורות משקלם גדול, ומותיר שורות ספורות בלבד ללב החיבור. בנוסף, בחיבור קצר כל כך משקלן של טעויות מקריות גבוה. גם הבדיקות הסטטיסטיות לא הראו עדיפות של צמדי הגריינים לעומת הגריינים הקצרים הבודדים.

ויתור על מטלות שילוב הקטעים

מטלת שילוב הקטעים היא מטלה מורכבת, ובשל כך צפוי היה שתבחין היטב בין נבחנים בעלי יכולת גבוהה לנבחנים חלשים. ברם, בניסוי נמצא כי מתאמי מטלות שילוב הקטעים עם הציון הפסיכומטרי הכללי ועם הציון בתחום המילולי היו דווקא נמוכים באופן ניכר ממתאמי המטלות האחרות. יתרה מזו, למורכבות המטלה יש גם צדדים שליליים: פיתוח מטלות מדגם זה מצריך זמן ארוך בהרבה מהזמן שנדרש לפיתוח מטלות פשוטות יותר, המבוססות על קטע אחד. כמו כן, המורכבות עלולה לגרום להחרפת תחושת לחץ הזמן של הנבחנים: אמנם לא נמצאו באיכות החיבורים ובסגנונם עדויות בולטות למצוקת זמן שהנבחנים היו נתונים בה, אך מסקירה של המשובים שמולאו בתום הניסוי עולה שהנבחנים במטלות מדגם זה חשו לחוצים יותר מהנבחנים האחרים. נראה שמורכבותה של המטלה מתאימה יותר לאוכלוסיית נבחנים חזקה יותר והטרוגנית פחות מאוכלוסיית נבחני מכפ"ל.

ויתור על הגריינים הארוכים בצורתם הראשונית

הגריינים הארוכים שנבדקו בניסוי הם למעשה גריינים למטלות קריאה-כתיבה. למטלות אלה שני יתרונות חשובים: רמת האוטנטיות שלהן גבוהה יחסית, שכן הן דומות למטלות אקדמיות, והן מצמצמות את ההסתמכות על ידע קודם. עם זאת, המשקל של יכולת הבנת הנקרא בהן רב, והדבר מגביל את היכולת להעריך את יכולת הכתיבה בנפרד. בניסוי נמצא כי חיבורים רבים שנכתבו בתגובה לגריינים מסוג זה התבססו במידה רבה מדי על הטקסט שבגריין, הן בממד התוכן הן בממד הלשון. בחיבורים של כ-30 שורות שיעורו של טקסט הגריין בחיבור גבוה מאוד. עוד נמצא כי אמנם השימוש בקטעי מקור בתור גריינים למטלת כתיבה תורם לאוטנטיות, אך הוא מגביל במידה ניכרת את היכולת לספק מידע.

צמצום השימוש בגריינים הקצרים בצורתם הנוכחית, ויצירת דגם המשלב בין הדגם הקצר ובין

הדגם הארוך

לדגם הקצר יתרונות בלתי מבוטלים: הוא מערב במידה המועטה ביותר את יכולת הבנת הנקרא, ולפיכך הוא הקרוב ביותר לאידאל של מטלת כתיבה "טהורה". כמו כן, פיתוח מטלות מדגם זה פשוט וזול יחסית, והן קלות לתרגום. משקלם של יתרונות אלה גדול נוכח העובדה שיידרוש עשרות מטלות תפעוליות בשנה. עם זאת, לדגם זה שני חסרונות בולטים: ראשית, הוא אינו אותנטי – כתיבה עיונית אינה נעשית יש מאין, אלא מבוססת כמעט תמיד על קריאה ממוקדת; ושנית, הוא אינו מספק הקשר לכתיבה, והנבחנים עשויים להתקשות לגבש דעה בנושא תלוש מהקשר. ביטוי לקושי הזה נמצא הן בקריאת החיבורים בניסוי והן במשובי הנבחנים – כ-6% מהנבחנים שנבחנו במטלות הגריין הקצר התלוננו במשוב על היעדר בחירה, לעומת כ-1.5% מהנבחנים במטלות מן הסוגים האחרים.

לאור דברים אלה התקבלו שתי החלטות הנוגעות להמשך הפיתוח של דגם זה: הראשונה, לבדוק את האפשרות להשתמש בגריינים קצרים במקרים שהיעדר הקשר אינו מפריע לכתיבה.

השנייה, ליצור דגם המשלב בין הדגם הקצר ובין הדגם הארוך. נקבע כי אורך הגריין והמטלה יהיה עד 10 שורות – אורך מספיק כדי להעלות את רמת האוטנטיות של המטלות הקצרות, לספק הקשר לכתיבה ולצמצם הסתמכות על ידע כללי, אך קצר מזה של המטלות הארוכות. בניסוח המטלה תהיה הקפדה על שפה פשוטה ובהירה, גם במחיר ויתור-מה על אותנטיות (כלומר על שימוש בקטעי מקור), כדי לצמצם ככל האפשר תלות ביכולת הבנת הנקרא ולאפשר הצגת הקשר רחב יותר בהיקף טקסט נתון.

חיזוק הזיקה בין המטלה ובין הגריין, וויתור על מטלות שהגריין מספק בהן רק קונטקסט

כמה מהגריינים שנבדקו בניסוי שיקפו פשרה בין הרצון לספק לנבחנים הקשר לבין הרצון לצמצם ככל האפשר את התלות ביכולת הבנת הנקרא: הטקסט בגריינים האלה אמור לתת לנבחן "השראה", אך הנבחן אינו חייב להבינו לעומק, ואף יכול להצליח במטלה ללא קריאתו (לדוגמה, מטלת מגבלותיו של המחשב, שהוצגה לעיל). שימוש בגריינים שכמעט שאינם נחוצים לביצוע המטלה בעייתי מכמה סיבות: ראשית, תפקידו של הגריין בכתיבת החיבור אינו ברור, ועמימות זו עלולה לפגוע בנבחנים; שנית, השימוש בגריינים כאלה מביא לבזבוז הזמן המוקדש

לקריאת הטקסט, שאינו נחוץ לכתיבה; ולבסוף, אי-בהירות בנוגע לטיב הקשר בין הגריין ובין המטלה מקשה על פיתוח מטלות מסוג זה. לפיכך הוחלט להשתמש רק בגריינים שזיקתם למטלה גבוהה.

ניסוח עקרונות בסיסיים בחיבור מטלות

נוסחו כמה עקרונות שיהיו הבסיס לכתיבת מטלות בהמשך תהליך הפיתוח:

- עקרון הקיצור – כדי למזער את התלות בהבנת הנקרא הוחלט שהגריין יהיה קצר יחסית (עד 10 שורות, כולל המטלה), אך לא קצר מכדי לספק הקשר לכתיבה.
- עקרון הפשטות – מאותו הטעם, הוחלט שהגריין יהיה כתוב בלשון פשוטה ובהירה ככל האפשר, אגב העדפה של טקסטים עיוניים-מיידעים על פני טקסטים שיש בהם ממד רטורי או אסתטי.
- עקרון ההקשר – כאשר נושא המטלה זר לנבחנים רבים, יש לספק הקשר. במטלות שבהן נראה שחיבור טוב חייב לגעת בהקשר המובא במטלה, יש לדרוש מפורשות מהנבחנים להתייחס להקשר.
- עקרון ההתרחקות מהבנאלי – כדי להימנע ממצב ששינון של כמה חיבורים גנריים יספיק כדי להצליח במטלת הכתיבה, יש לבחור בנושאים בנאליים כמה שפחות.

ניסוי שני

בעקבות תוצאות הניסוי הראשון ומסקנותיו, צוות החשיבה המילולית החל לפתח ולחבר מטלות כתיבה, כדי לגבש מסקנות מבוססות הנוגעות למאפייני מטלות הכתיבה שישמשו בניסוי השני, ולהתחיל ביצירת בסיס לבנק מטלות תפעוליות.

בתהליך הכתיבה נמצא שרבות מהמטלות הנכתבות דומות זו לזו באופי הגריין שהן מציגות – דילמה אתית שבמרכזה התנגשות בין שני ערכים – ובאופי המטלה הנגזרת ממנו, כתיבה טיעונית ובה ננקטת עמדה מנומקת בדילמה המוצגת. לדוגמה:

שכרו של דיוויד לטרמן, מנחה תכנית אירוח מפורסמת בטלוויזיה בארצות הברית, גבוה פי 700 משכרו של פועל במה העובד בתכניתו.

לדעתכם, האם יש הצדקה לפער השכר הגדול בין השניים? נמקו את עמדתכם.

מצב זה עלול להיות בעייתי משתי סיבות עיקריות: הראשונה – חשש למיצוי מהיר של מאגר המטלות הפוטנציאלי. השנייה – הדמיון בין המטלות עלול להביא להכנה מוקדמת של חיבורים, שכן יהיה קל יחסית להתאימם לנושא הספציפי המופיע בבחינה.

כדי להרחיב את מאגר המטלות הפוטנציאלי, הוחלט לבדוק גריינים ומטלות בעלי אופי שונה. את אלה אפשר לחלק לשלוש קבוצות עיקריות:

1. מטלות לימודיות – מטלות שבגריין שלהן מוצג מידע אקדמי ברמה בסיסית, ובמטלה נדרשת התייחסות ישירה למידע זה.

■ דוגמה

בשנת 2004 הזדעזע העולם מחשיפת מסכת התעללויות קשות שעברו על אסירים עיראקים בכלא אבו עריב שבעיראק מידי סוהריהם, חיילים אמריקנים. כמה פוליטיקאים אמריקנים טענו כי אירועי אבו עריב התרחשו בשל כמה חיילים שמאפייני אישיותם בעייתיים, כמו נטייה קיצונית לאלימות. בעקבות טענות אלו כתב הפסיכולוג פיליפ זימברדו מאמר, ובו טען כי לעתים קרובות הנסיבות הן שאחראיות להתנהגות האלימה וגורמות גם לבני אדם רגילים לנהוג באכזריות. הוא הסתמך על ניסוי שעשה בשנת 1971. בניסוי הוא חילק קבוצת סטודנטים באופן אקראי ל"אסירים" שנכלאו במרתף, ול"סוהרים". ה"סוהרים" צידו במדים ובאלות והונחו לשמור על הסדר ב"כלא". אחרי זמן קצר נהגו ה"סוהרים" ב"אסירים" באלימות קשה כל כך עד שהוחלט להפסיק את הניסוי לפני המועד המתוכנן. לדברי זימברדו, הצבת החיילים האמריקנים בתפקיד של סוהרים היא שגרמה להם להתעלל באסירים, ממש כמו בניסוי שעשה.

על סמך טענתו של זימברדו, האם יש מקום, לדעתכם, להקל בעונשם של החיילים המתעללים, ואולי אף לזכותם מאשמת התעללות? נמקו את דבריכם תוך התייחסות לניסוי.

2. מטלות תלויות קונטקסט – מטלות שנדרשת בהן כתיבה טיעונית בנוגע לדילמה אתית שנטועה בקונטקסט ספציפי.

■ דוגמה

בשנת 1979 הכריזה ממשלת סין על תכנית להגבלת הילודה. סין היא המדינה המאוכלסת ביותר בעולם, והתכנית נועדה למנוע פיצוץ אוכלוסין – מצב שבו יהיה מספרם של תושביה גדול כל כך עד כי משאבי המדינה לא יוכלו לספק את צרכי כולם. כחלק מן התכנית השקיעה הממשלה ממון רב בפעולות הסברה ובחלוקת אמצעי מניעה. כמו כן, הממשלה העניקה מענקים כספיים והטבות נוספות לזוגות שהתחייבו שלא להביא לעולם יותר מילד אחד. אחר כך הפעילה הממשלה סנקציות כלכליות על נשים בהריון שני כדי לגרום להן לבצע הפלה מלאכותית.

מה דעתכם על החלטתה של סין להתמודד עם סכנת פיצוץ האוכלוסין באמצעות תכנית להגבלת ילודה?

3. מטלות תיאוריות – מטלות שהנבחנים נדרשים לכתוב בהן חיבור תיאורי.

■ דוגמה

הטלוויזיה נחשבת המצאה שהשפעתה על החברה המודרנית עצומה. השפעתה נוגעת לפרט ולחברה בכמה היבטים, למשל תפיסת העולם של הפרט, כישוריו, התנהגותו, יחסיו עם סביבתו ואף דמותה של החברה כולה.

תארו כיצד הטלוויזיה משפיעה על אחד מן ההיבטים שלעיל או על כמה מהם.

תוצאות הניסוי השני

הניסוי השני נערך במתכונת דומה לניסוי הראשון. בניסוי נבדקו שמונה מטלות כתיבה – שש מטלות שנכתבו בידי צוות הפיתוח בעברית ובהן ייצוג לשלוש הקבוצות שהוזכרו, ושתי מטלות נוספות שפותחו בידי צוות הפיתוח בערבית ותורגמו לעברית. בניסוי השתתפו 929 נבחנים בשפה העברית (כ-115 נבדקים למטלה) ו-325 נבדקים בשפה הערבית (בערבית הועברו 4 מטלות בלבד – כ-80 נבדקים למטלה). גם בניסוי זה המשתתפים היו נבחנים שעמדו להיבחן במכפ"ל כשבוע לאחר הניסוי.

טבלה מספר 3 – מתאמים בין הערכות שני המעריכים בעברית (ציון כללי במטלה), $n=929$

מתאם בין הערכות המעריכים	
0.59	מטלות מקוריות בעברית (6 מטלות, $n=716$)
0.42	מטלות מתורגמות (2 מטלות, $n=213$)
0.56	כל המטלות יחד (8 מטלות, $n=929$)

(תוצאות הניסוי הראשון היו כזכור: גריין קצר – 0.48, צמדים – 0.51, גריין ארוך – 0.61, שילוב קטעים – 0.53).

בכל המדדים שנבדקו נמצאו פערים בין המטלות המקוריות ובין המתורגמות, לטובת המקוריות. הסיבה לכך אינה ברורה, אבל מהבדיקה המצומצמת הזאת עולה שכדאי לשקול הימנעות משימוש במטלות מתורגמות (ר' גם דיון בנושא זה בהמשך).

טבלה מספר 4 – מתאמים בין ציוני מטלת הכתיבה ובין ציוני האמת בבחינה בתחום המילולי (VE) ובבחינה כולה (TOT) בעברית, $n=929$

	תוכן	לשון	כללי	
מטלות מקוריות בעברית (6 מטלות, $n=716$)	VE	0.46	0.55	0.54
	TOT	0.46	0.52	0.52
מטלות מתורגמות (2 מטלות, $n=213$)	VE	0.29	0.40	0.38
	TOT	0.27	0.36	0.35
כל המטלות (8 מטלות, $n=929$)	VE	0.43	0.52	0.51
	TOT	0.43	0.49	0.49

מתאמי הציונים במטלות עם הציונים בתחום המילולי ובבחינה כולה עלו לעומת הניסוי הראשון (בניסוי הראשון הן היו כזכור : $VE = 0.18-0.44$, $TOT = 0.26-0.44$). גם בניסוי השני נמצא יתרון עקיב לממד הלשון, אם כי יתרון זה היה נמוך מהיתרון שנמצא בניסוי הראשון. הסיבות האפשריות לכך נסקרו בתיאור הניסוי הראשון. בניסוי הראשון הפער הגדול ביותר בין הלשון לתוכן היה במטלות צמדי הגריינים הקצרים – כאמור, כנראה מפני שקשה לפתח תוכן של ממש במסגרת זמן של 15 דקות. אי-שימוש בגריינים מסוג זה בניסוי השני תרם לצמצום הפער בין שני הממדים.

תרגום מטלות

בערבית הועברו כאמור ארבע מטלות בלבד : שתי מטלות שנכתבו בערבית במקור ושתיים שתורגמו מעברית. מתאמי הציונים במטלות המקוריות עם ציוני האמת בתחום המילולי היו 0.68 ו-0.45, ומתאמי הציונים במטלות המתורגמות עם ציונים אלה היו 0.43 ו-0.44. המתאמים בין ציוני המטלות שנכתבו במקור בערבית ובין הביצוע בתחום המילולי בבחינה גבוהים מהמתאמים של ציוני אותן מטלות, כשתורגמו לעברית, עם הציון המילולי בבחינה (0.57 ו-0.38 בהתאמה). באופן דומה, גם המתאמים בין ציוני המטלות המתורגמות מעברית ובין הביצוע בתחום המילולי בבחינה נמוכים לעומת אותן מטלות כשהועברו בשפת המקור (0.43 ו-0.54 בהתאמה). נתונים אלה מצטרפים לפערים שנמצאו בין מתאמי הציונים של המטלות המקוריות למתורגמות בשתי השפות, ומעידים כי מוטב להשתמש במטלות הכתובות בשפת המקור ולא במטלות מתורגמות.

טבלה מספר 5 – המתאמים בין ציוני מטלת הכתיבה בעברית ובין ציוני האמת בבחינה כולה / בתחום המילולי

		ציון פסיכומטרי כללי			ציוני התחום המילולי		
שם המטלה	סוג הגריין והמטלה	תוכן	לשון	כללי	תוכן	לשון	כללי
השפעת הטלויזיה על חיינו (n=155)	- גריין קצר - מטלת כתיבה תיאורית	0.37	0.44	0.43	0.43	0.53	0.51
החמרת תקנות העישון במקומות ציבוריים (n=129)	- גריין המבוסס בחלקו על קטע מקור אותנטי [מאמר עמדה בעיתון] - מטלת כתיבה טיעונית	0.57	0.55	0.59	0.61	0.59	0.63
מדוע חשוב ללמוד היסטוריה בבתי הספר (n=96)	- גריין קצר - מטלת כתיבה טיעונית חד-כיוונית ["מדוע"]	0.44	0.52	0.50	0.44	0.52	0.50

0.60	0.60	0.49	0.60	0.58	0.50	- גריין לימודי [הצגת טקסט אקדמי חדש] - מטלת כתיבה טיעונית	ניסויי זימברדו – מנגנונים פסיכולוגיים של התעללות (n=115)
0.58	0.61	0.46	0.59	0.59	0.49	- גריין קצר המציג דילמה אתית בקונטקסט נתון - מטלת כתיבה טיעונית	האם פערי שכר קיצוניים מוצדקים (n=104)
0.41	0.44	0.33	0.45	0.46	0.40	- דילמה אתית המוצגת בקונטקסט נתון - מטלת כתיבה טיעונית	הצעה להוריד את גיל המינימום לבחירה לכנסת (n=117)
0.37	0.40	0.28	0.34	0.36	0.27	מתורגם מערבית	האם כדאי לשדר פחות ריאליטי בטלוויזיה (n=112)
0.40	0.41	0.31	0.37	0.39	0.28	מתורגם מערבית	מה גורם לריבוי תאונות הדרכים (n=101)

בשתיים משלוש המטלות שהמתאמים שלהן עם הבחינה היו הגבוהים ביותר (החמרת תקנות העישון, זימברדו), ייתכן שהמתאמים מוטים כלפי מעלה בגלל תלות גדולה שלהן ביכולת הבנת הנקרא, שמוערכת גם בפריטי הבררה של הבחינה.

מסקנות מן הניסוי השני בנוגע להמשך הפיתוח וליצירת בנק מטלות תפעוליות

מעבודת פיתוח המטלות לקראת הניסוי, מנתוני הניסוי ומהמשובים עליו (של נבחנים, בוחנים ומעריכי חיבורים) עלו כמה מסקנות הנוגעות להמשך תהליך הפיתוח וליצירת בסיס לבנק המטלות התפעוליות:

גריינים

1. התמקדות בפיתוח מטלות המבוססות על גריינים "לימודיים" (המציגים חומר אקדמי חדש) ועל גריינים שהדילמות בהם תלויות קונטקסט – שני סוגי הגריינים האלה פותחו כמסקנה מעבודת הפיתוח לקראת הניסוי השני ומתוצאותיו. הם נועדו לתת מענה לבעיות שאפיינו את הגריינים הארוכים (תלות גדולה ביכולת הבנת הנקרא, חיבורים שחלקים גדולים מהם לקוחים מהגריין), ולא להשאיר את הגריינים הקצרים (היעדר הקשר, הסתמכות על ידע קודם, מאגר נושאים מוגבל). במהלך תהליך הפיתוח התחזקה בקרב צוות הפיתוח הדעה ששני הסוגים נותנים מענה לבעיות אלה, ושהם שקולים זה לזה בהיקפם ובדרישות שהם מציבים בפני הנבחנים, כך שניתן להשתמש בשניהם מבלי לפגוע בסטנדרטיות של הבחינה.

■ דוגמה למטלה לימודית

"ניסוי בנדורה" הוא ניסוי מפורסם שערך הפסיכולוג אלברט בנדורה בשנת 1961. בחלקו הראשון של הניסוי חילק בנדורה ילדים לשתי קבוצות. קבוצה אחת צפתה באדם בוגר מכה בובה ומקלל אותה, והקבוצה האחרת צפתה באדם בוגר מסדר צעצועים בשלווה. בשלב השני של הניסוי הכניס בנדורה את שתי קבוצות הילדים לחדר אחד ובו צעצועים ובובות. אחר כך צוות החוקרים השתמש בשיטות שונות להרגיז את הילדים ורשם את תגובותיהם. נמצא כי רמת התוקפנות של הילדים מהקבוצה שצפתה באדם האלים הייתה גבוהה בהרבה מרמת התוקפנות של הילדים שהשתייכו לקבוצה האחרת. חלקם אף חזרו על מעשיו, הכו בבובות והשמיעו צעקות כמוהו. בנדורה הסיק מכך כי ילדים שנחשפים לאלימות ירבו לנהוג באלימות.

ממצאי הניסוי הביאו רבים לטעון כי יש להתחשב בילדותם של פושעים אלימים בעת העמדתם לדין. לטענתם, הניסוי מראה שפושעים שנחשפו בילדותם לאלימות הושפעו מכך ולכן הפכו לאלימים בעצמם. הביעו דעתכם על טענה זו, תוך התייחסות לניסוי בנדורה.

■ דוגמה למטלה דילמה תלוית קונטקסט

ב-1982 הגו הקרימינולוגים האמריקנים ווילסון וקלינג תאוריה, ולפיה במקומות שרווחים בהם פשעים קלים, גם פשעים חמורים ייעשו רווחים במוקדם או במאוחר, שכן במקומות האלה נוצרת אווירה של אי-סדר ואימה, אווירה שהיא קרקע פורייה להתפתחות פשיעה חמורה. לכן, לדעתם, המלחמה בפשע מחייבת חמרת הענישה על עברות קלות. תאוריה זו הייתה הבסיס לניסוחו של "חוק שלוש הפסילות" ולהחלתו בכמה מהמדינות בארצות הברית. לפי חוק זה, על כל עברייני שהורשע בביצוע שלושה פשעים שונים, גם אם כולם נחשבים לעברות רכוש קלות בלבד (גניבת חולצה מחנות בגדים למשל), מוטל עונש כבד מאוד: ממאסר של 25 שנים עד מאסר עולם, ללא אפשרות לשחרור מוקדם. מאז הוחל החוק חלה ירידה חדה בפשיעה – קלה וחמורה כאחת – באזורים שהוא הוחל בהם. ואולם רבים טוענים כי חוק זה פסול, בעיקר מכיוון שענישה שאינה תואמת את חומרת העברה אינה צודקת ואינה מוסרית.

האם אתם תומכים ב"חוק שלוש הפסילות"? נמקו.

2. ויתור על גריינים קצרים – משני הניסויים עולה כי גריינים אלה בעייתיים מפני שאינם מספקים הקשר ומחייבים הסתמכות על ידע קודם. לבעיות אלה, שנדונו בעקבות הניסוי הראשון, נוספת העובדה שהכללת גריינים קצרים עלולה לפגוע בסטנדרטיות של הבחינה, שכן הם אינם שקולים, באורכם וביכולות שהם דורשים מהנבחן, לגריינים הלימודיים ולגריינים המציגים דילמות תלויות קונטקסט.

סוגת הכתיבה שתידרש מהנבחנים

התמקדות במטלות הדורשות כתיבה עיונית-טיעונית וויתור על מטלות הדורשות כתיבה תיאורית – כאמור, מכיוון שכתובה עיונית-טיעונית היא הסוגה הנפוצה באקדמיה, רצוי שמטלת כתיבה הבוחנת פוטנציאל כתיבה אקדמית תתבסס על אותה סוגה. עם זאת, בניסוי השני נבדקה גם האפשרות להכללת מטלות כתיבה תיאורית. בניסוי נבדקה מטלה אחת כזאת ("השפעת הטלוויזיה על חיינו"), ובתהליך הפיתוח נכתבו כמה מטלות תיאוריות נוספות. בעקבות תהליך פיתוח המטלות וסקירת החיבורים שנכתבו על ידי הנבחנים הוחלט לא להמשיך בפיתוח מטלות תיאוריות, וזאת מכמה טעמים: ראשית, אף שמטלות כאלה קיימות גם

באקדמיה, הרי שהן דורשות יכולות נמוכות יותר מאלה הנדרשות לכתיבה טיעונית. זאת ועוד, כאשר המטלות האלה מופיעות בהקשר אקדמי, בדרך כלל הן דורשות הסתמכות על ידע שנצבר בזמן הלימודים. בסיטואציה של בחינת כניסה לאקדמיה משמעות הדבר היא בדרך כלל הסתמכות על ידע קודם של הנבחן, שרצוי לצמצמה ככל האפשר. בנוסף לכל אלה, שימוש במטלות תיאורית לצד מטלות כתיבה טיעונית יפגע בסטנדרטיות של המבחן. ולבסוף, ממצאי מחקר מעידים כי במטלות כתיבה טיעונית החיבורים מאופיינים ברמה לשונית גבוהה יותר מחיבורים שנכתבו בתגובה למטלות כתיבה תיאורית (Crowhurst & Piche, 1979), דבר שמאפשר טווח ביצוע גדול יותר ויכולת הבחנה טובה יותר בין הנבחנים.

מסקנות נוספות

1. המשך בדיקה של מטלות מסוגים חדשים – הוחלט להמשיך לבדוק ולפתח סוגים חדשים של מטלות.

▪ דוגמה – מטלה לימודית-יישומית

חוקי היסוד בישראל קובעים כי אסור למדינה לפגוע בזכויות אדם: אסור למשל לפגוע בגופו של אדם או בחייו, בכבודו, בקניינו, בחופש התנועה שלו ובפרטיותו. עם זאת, יש בחוקי היסוד סעיף המאפשר למדינה לפגוע בזכויות אדם במקרים חריגים. הסעיף מכונה **פסקת ההגבלה**, והוא קובע שהפגיעה בזכויות צריכה להיות **לתכלית ראויה**, וכן שעליה להיות **מידתית** – כלומר שאי אפשר להשיג את אותה התכלית באמצעות פגיעה קטנה יותר בזכויות אדם. כך למשל גביית מסים אמנם פוגעת בקניינם של האזרחים, אך היא מיועדת לתכלית ראויה – יצירת מערכות בריאות, חינוך, ביטחון וכדומה. תכלית זו לא תושג בלי גביית מס מהאזרחים, ולפיכך הטלת מסים היא פגיעה מידתית בזכות הקניין.

לאחרונה עלתה הצעה להעלות את הגיל שבו אפשר לקבל רישיון נהיגה מ-17 ל-19. ההצעה עלתה לאחר שנמצא כי שיעור הנפצעים קשה בתאונות דרכים בקרב הנהגים עד גיל 19 גבוה ב-30% מזה שבקרב הנהגים בגילאי 19-24.

האם החוק המוצע עומד לדעתכם בדרישות פסקת ההגבלה? נמקו.

2. שימוש במטלות מתורגמות – נוכח הנתונים שעלו מהעברת מטלות מתורגמות מעברית לערבית ומערבית לעברית, הוחלט להימנע משימוש במטלות מתורגמות מערבית בבחינה העברית. כרגע לא ניתן לפתח מספיק מטלות מקוריות בערבית עבור האוכלוסייה הנבחרת בשפה זו, אך השאיפה היא להקטין ככל האפשר את שיעור המטלות המתורגמות מעברית בשנתיים הקרובות.

3. בניית בנק מטלות תפעוליות – עם סיכום העבודה על הניסוי השני ותוצאותיו, החל צוות החשיבה המילולית לעבוד על בניית בנק מטלות תפעוליות. היעד הוא לפתח בשנתיים הקרובות 160 מטלות.

חלק ג – הערכת החיבורים ונקידתם

על הערכת המטלה להיות מהימנה ותקפה ככל האפשר. מהימנות גבוהה תבוא לידי ביטוי במידת הסכמה רבה בין המעריכים. התוקף יהיה גבוה ככל שהציון ישקף את יכולת הכתיבה, וככזה אפשר להניח שיהיה מתואם באופן חיובי עם משתנים שנמדדות בהם יכולות מילוליות. כדי להשיג מטרות אלה יש צורך במעריך הערכה ונקידה לחיבורים שיהיה תקף, מובנה ואובייקטיבי. על המעריך לכלול שיטת הערכה – הוליסטית או אנליטית; שיטת נקידה – סולם הציונים המשמש להערכה; הליך הערכה – קביעת מספר המעריכים ואופן חישוב הציון; מחוון – מדריך להערכת חיבורים; וכן מנגנון איתור מעריכים, הכשרה שלהם ומעקב אחרי עבודתם. אופי המעריך נגזר במידה רבה מהשיטה שבוחרים להעריך בה את החיבורים: הערכה הוליסטית או הערכה אנליטית.

שיטות ההערכה הקיימות

הערכה הוליסטית

שיטת ההערכה ההוליסטית היא הנפוצה ביותר בפרויקטים רחבי היקף של הערכת יכולות כתיבה, ומשמשת בבחינות המיון המרכזיות בארצות הברית (GRE, SAT). ההערכה ההוליסטית מבוססת על תפיסת הכתיבה כיכולת אחת, שאי אפשר, או לא כדאי, להעריכה בדרך של פירוקה להיבטים שונים. לפיכך, בהערכה בגישה זו החיבור מקבל מהמעריך ציון אחד. מן הגישה ההוליסטית נגזרות שיטות הנבדלות אלה מאלה בדרך קביעת הציון: יש שיטות המסתמכות על הרושם הכללי הראשוני שהמעריך מקבל מהטקסט שקרא, ויש שיטות המבוססות על מעין שקלול של איכות הטקסט בכמה ממדים קבועים.

להערכה ההוליסטית יתרונות פרקטיים ניכרים: ההערכה בה אינטואיטיבית וקלה יחסית, ומכיוון שנדרשת בה קריאה אחת בלבד של החיבור, היא גם מהירה. על כן היא גם זולה יחסית לחלופה האנליטית, שתוצג להלן.

עם זאת, להערכה ההוליסטית גם כמה חסרונות בולטים. הנחיה להערכה הוליסטית המכוונת את המעריך להתייחסות לכמה ממדים עלולה לגרום לכל מעריך לסכם באופן סובייקטיבי את הערכותיו לממדים אלה. משמעות הדבר היא שמעריכים שונים עלולים להקצות משקלות שונים לממדים השונים – ויתרה מזאת, אפילו אותו מעריך עלול להקצות משקלות שונים לממדים השונים בנקודות זמן שונות או ביחס לחיבורים שונים. הדבר עלול להביא לרעש במדידה ולפגוע במהימנות ההערכות. הערכה הוליסטית שאינה מבוססת על כמה ממדים קבועים בעייתית עוד יותר, שכן הערכות כאלה תלויות בפרשנות ובנטייה האישית של כל מעריך, והתוצאה היא שונות בהערכות שחלק לא מבוטל ממנה מקורו במעריכים. חיסרון נוסף של ההערכה ההוליסטית הוא שקשה להגיע להסכמה בדבר הציון ההוליסטי שיש לתת לחיבור טוב מאוד בממד מסוים וגרוע באחר, מה שמקשה את ההשוואה בין ציונים של נבחנים שונים באותו נושא, כמו גם בין ציונים של אותם נבחנים על חיבורים בנושאים שונים.

הערכה אנליטית

ההערכה האנליטית מבוססת על התפיסה שיכולת הכתיבה מורכבת מכמה מרכיבים נפרדים, ולפיכך יש להעריך את המרכיבים אלה בנפרד ובאופן בלתי תלוי. נקידה אנליטית מבוססת על זיהוי הממדים החשובים בכתיבה, אופרציונליזציה שלהם ומתן משקלות יחסיים לכל אחד מהם. על הממדים לעמוד בשתי דרישות מנוגדות: עליהם להיות מדויקים ככל האפשר, כדי למזער את המרכיב הסובייקטיבי בהערכות השופטים ולהגיע למהימנות גבוהה, ובה בעת כלליים מספיק כדי שיתאימו למטלות כתיבה שונות. אפשר להעריך את הממדים באופן כמותי או באופן איכותי. הערכה כמותית מבוססת על ספירה של מרכיבים שונים בטקסט המוערך (טעויות דקדוק, טעויות איות, מספר מילים במשפט וכדומה), והיא אובייקטיבית באופייה, ואילו הערכה איכותית היא הערכה סובייקטיבית יותר של איכות הממדים השונים.

להערכה האנליטית כמה יתרונות חשובים: ראשית, קיומם של קריטריונים ברורים לנקידה מביא לתהליך הערכה ממוקד, קפדני ואובייקטיבי יותר, ועוזר לצמצם השפעה של גורמים לא רלוונטיים שמקורם במעריכים. שנית, שלא כמו בהערכה הוליסטית המבוססת על כמה ממדים, בהערכה אנליטית צירוף המידע מהממדים השונים נעשה בדיוק ובעקבות. שלישית, התוקף הנראה של שיטות אנליטיות גבוה יותר: הערכה אנליטית נראית אובייקטיבית והוגנת יותר מהערכה הוליסטית, ודומה שהיא מאפשרת להשוות ביתר קלות בין ציונים של חיבורים שנכתבו בתגובה למטלות שונות.

לשיטה גם כמה חסרונות: פירוק הכתיבה לממדים הוא מלאכותי, ועלול להביא להערכה המעניקה משקל רב להיבטים מסוימים של יכולת הכתיבה אגב התעלמות מכך שיכולת הכתיבה היא יותר מסך רכיביה, וממילא אין אפשרות מעשית לבחון כל רכיב ורכיב שלה. מעבר לכך, הערכה אנליטית אורכת זמן רב יחסית – גם כשממדי ההערכה מעטים – מכיוון שיש צורך לקרוא כל חיבור כמה פעמים, כמספר ממדי ההערכה.

סוגי מחוונים

במקביל לפיתוח המטלות לקראת הניסוי הראשון נערכו בדיקות ראשוניות של שני סוגי מחוונים, הוליסטי ואנליטי. לבדיקות אלה היו שתי מטרות: יצירת כלי הערכה שיסייע בפיתוח המטלות בשלבי הפיתוח הראשוניים, ופיתוח מחוון שישמש להערכת מטלת הכתיבה. בשלב הראשון נבחנו המחוון של ה-SAT והמחוון שפותח במאל"ו בשנים 1995-1997, המייצגים שתי גישות שונות בנוגע לדרך ההערכה, לאופן הנחיית המעריך ולדרגת החופש שהוא מקבל, לתכונות המוערכות ולאופן ההצגה הגרפית שלהן. מחוונים אלו שימשו בסיס להכרעה בסוגיות המרכזיות שעמדו לפני צוות הפיתוח בעת בניית המחוון.

מחווה הוליסטית

מדריכי הנקידה של ה-SAT ושל מטלת הצגת העמדה ב-GRE דומים במרבית המאפיינים שלהם (מטלת הכתיבה הנוספת ב-GRE – ניתוח ביקורתי של טיעונים – שונה באופייה מהמטלות שפותחו במאל"ו, ונגזרות ממנה הנחיות נקידה שונות, רלוונטיות פחות לצורכי מכפ"ל). לצורך הבדיקות הראשוניות השתמשנו במחווה ה-SAT:

The New SAT Scoring Guide

SCORE OF 6 An essay in this category is <i>outstanding</i> , demonstrating <i>clear and consistent mastery</i> , although it may have a few minor errors. A typical essay - effectively and insightfully develops a point of view on the issue and demonstrates outstanding critical thinking, using clearly appropriate examples, reasons, and other evidence to support its position - is well organized and clearly focused, demonstrating clear coherence and smooth progression of ideas - exhibits skillful use of language, using a varied, accurate, and apt vocabulary - demonstrates meaningful variety in sentence structure - is free of most errors in grammar, usage, and mechanics	SCORE OF 5 An essay in this category is <i>effective</i> , demonstrating <i>reasonably consistent mastery</i> , although it will have occasional errors or lapses in quality. A typical essay - effectively develops a point of view on the issue and demonstrates strong critical thinking, generally using appropriate examples, reasons, and other evidence to support its position - is well organized and focused, demonstrating coherence and progression of ideas - exhibits facility in the use of language, using appropriate vocabulary - demonstrates variety in sentence structure - is generally free of most errors in grammar, usage, and mechanics	SCORE OF 4 An essay in this category is <i>competent</i> , demonstrating <i>adequate mastery</i> , although it will have lapses in quality. A typical essay - develops a point of view on the issue and demonstrates competent critical thinking, using adequate examples, reasons, and other evidence to support its position - is generally organized and focused, demonstrating some coherence and progression of ideas - exhibits adequate but inconsistent facility in the use of language, using generally appropriate vocabulary - demonstrates some variety in sentence structure - has some errors in grammar, usage, and mechanics
SCORE OF 3 An essay in this category is <i>inadequate</i> , but demonstrates <i>developing mastery</i> , and is marked by ONE OR MORE of the following weaknesses: - develops a point of view on the issue, demonstrating some critical thinking, but may do so inconsistently or use inadequate examples, reasons, or other evidence to support its position - is limited in its organization or focus, or may demonstrate some lapses in coherence or progression of ideas - displays developing facility in the use of language, but sometimes uses weak vocabulary or inappropriate word choice - lacks variety or demonstrates problems in sentence structure - contains an accumulation of errors in grammar, usage, and mechanics	SCORE OF 2 An essay in this category is <i>seriously limited</i> , demonstrating <i>little mastery</i> , and is flawed by ONE OR MORE of the following weaknesses: - develops a point of view on the issue that is vague or seriously limited, demonstrating weak critical thinking, providing inappropriate or insufficient examples, reasons, or other evidence to support its position - is poorly organized and/or focused, or demonstrates serious problems with coherence or progression of ideas - displays very little facility in the use of language, using very limited vocabulary or incorrect word choice - demonstrates frequent problems in sentence structure - contains errors in grammar, usage, and mechanics so serious that meaning is somewhat obscured	SCORE OF 1 An essay in this category is <i>fundamentally lacking</i> , demonstrating <i>very little or no mastery</i> , and is severely flawed by ONE OR MORE of the following weaknesses: - develops no viable point of view on the issue, or provides little or no evidence to support its position - is disorganized or unfocused, resulting in a disjointed or incoherent essay - displays fundamental errors in vocabulary - demonstrates severe flaws in sentence structure - contains pervasive errors in grammar, usage, or mechanics that persistently interfere with meaning

Essays not written on the essay assignment will receive a score of zero.

כפי שאפשר להתרשם, ההערכה של בחינות ה-SAT הוליסטית. בנוסף יש לה כמה מאפיינים נוספים שיש לזכור בעת ההתייחסות למחווה זו:

אופי המטלות

כזכור, המטלות הנכללות ב-SAT וב-GRE הן מטלות המבוססות על גריינים קצרים שאורכם כמה שורות בלבד, שהנבחן צריך לחוות את דעתו עליהם באופן מנומק. פירוט הממדים במחווה האמריקאית מתאים לאופי המטלות המובאות במבחנים אלו – גריינים קצרים, הדורשים הסתמכות על הידע הפרטי של המועמד ועל ניסיונו האישי, ולא על מידע המוצג בגריין עצמו.

שיטת הנקידה

המחווה המשמשים בבחינות המרכזיות בארצות הברית מספקים סולם הערכה שנע בין 1 (חלש מאוד) ל-6 (מעולה). שני מעריכים בלתי תלויים (בחלק מהבחינות אחד מהם הוא מחשב) נותנים כל אחד ציון בין 1 ל-6 במדרגות של חצאי נקודה. סכום הציונים הוא ציון הגלם של הנבחן, כלומר טווח הציונים הוא 2-12 (ולמעשה, מכיוון שהסולם מורכב מחצאי נקודות, הוא מקביל לסולם 4-24).

אופן הנחיית המעריך

מדריכי הנקידה של ה-SAT ושל ה-GRE מספקים פירוט של המאפיינים האיכותיים של כל אחת משש רמות הציון (שלא כמו המדריך שפותח במאל"י, המוצג בהמשך, אשר מציג רק את הממדים שעל פיהם יש להעריך את החיבורים, ואינו מפרט פירוט איכותי של כל אחת מרמות הציון בכל אחד מהממדים). בראשית התיאור של כל רמת ציון יש תיאור כללי של חיבור אופייני לציון זה (למשל: "ציון 6 – חיבור בקטגוריית ציון זו הוא חיבור מעולה, המפגין יכולת גבוהה, ברורה ועקיבה, אף שעשויות להופיע בו כמה טעויות זניחות"). אחרי התיאור הכללי מופיע פירוט המאפיינים האיכותיים: חמישה היבטים שהציון צריך לשקף, המבטאים את רמת התוכן, המבנה והלשון של תוצר הכתיבה האופייני לאותו ציון. לדוגמה (שוב, מציון 6) – "מבחן אופייני בציון זה:

- מפתח עמדה בנוגע לנושא המטלה ביעילות ואגב הפגנת תובנה, ומפגין חשיבה ביקורתית מעולה, אגב הבאת דוגמאות, סיבות וראיות אחרות שתומכות בעמדתו בבירור.

- מאורגן היטב וממוקד, מפגין לכידות ברורה ומעבר רציף בין רעיונות [...], "וכולי. יתרונו העיקרי של אופן ההנחיה הזה הוא שהוא חוסך מן המעריך את הצורך לקבוע עבור כל היבט או משתנה את הרמה שמתאימה לכל ציון. כל שנדרש מהמעריך הוא למצוא את התיאור המתאים ביותר לחיבור שקרא מבין תיאורי הציונים שבטבלה. בכך מופחת הסיכון לרעש שמקורו בפעולה נוספת, מלבד הערכת החיבור, שעל המעריך לבצע.

החיסרון העיקרי באופן ההנחיה הזה הוא שהוא מניח קיום של מתאם גבוה בין כל הממדים המפורטים, בשעה שלמעשה ייתכנו קומבינציות שונות של ביצוע בממדים השונים (למשל, רמה לשונית גבוהה מאוד לצד רמת תוכן נמוכה).

התכונות המוערכות

כאמור, ב-SAT מוצג בכל אחת מרמות הציון (1-6) פירוט הרמות של חמש תכונות המתאימות לאותה רמת ציון. התכונות מאפיינות את התוכן, המבנה והלשון שבחיבור אופייני לאותה רמה (אף שאין התייחסות מפורשת לחלוקה לשלושה ממדים אלה; החלוקה המוצגת להלן היא שלנו). נחזור לדוגמה של ציון 6:

תוכן: מפתח עמדה בנוגע לנושא המטלה ביעילות ואגב הפגנת תובנה, ומפגין חשיבה ביקורתית מעולה, אגב הבאת דוגמאות, סיבות וראיות אחרות שתומכות בעמדתו בבירור.

מבנה: מאורגן היטב וממוקד, מפגין לכידות ברורה ומעבר רציף בין רעיונות.

לשון: מפגין מיומנות בשפה, אגב שימוש באוצר מילים מגוון, מדויק והולם.

מראה גיוון בעל משמעות במבנה המשפטים.

כמעט ללא טעויות דקדוק, איות ושימוש.

הצגה גרפית

מדריכי הנקידה של ה-SAT ושל ה-GRE מופיעים בצורת טבלה. יתרונותיה של דרך הצגה זו הם נוחות השימוש ומהירות ההתמצאות, כמו גם קלות ההשוואה בין מאפייני רמות הציונים השונות.

בנוסף למדריכי הנקידה מקבלים המעריכים הנחיות מפורטות המתייחסות לגריין הספציפי לחיבורים שהם מעריכים.

מחווה אנליטי

לצורך הבדיקות הראשוניות השתמש צוות הפיתוח במחווה אנליטי דו-ממדי (תוכן ולשון), המבוסס על המחווה שפותח במאל"ו כחלק מתהליך הבנייה של מבחן הכתיבה בשנות התשעים. במסגרת תהליך זה נבדקו אמפירית שתי שיטות נקידה, הוליסטית ואנליטית (שכללה שלושה ממדים: תוכן, מבנה ולשון). השיטה ההוליסטית לא עמדה בציפיות מבחינת מהימנות ההערכה. הסיבה העיקרית לכך הייתה היעדר שיטה אובייקטיבית לשקלול הממדים השונים, מה שגרם לחוסר אחידות בין המעריכים בנוגע למשקל שניתן לכל ממד. נוסף על כך, המעריכים התקשו להשתחרר מ"אפקט ההילה" של ממד התוכן, ונתנו לו משקל רב יותר מלשאר הממדים. אשר לשיטת הנקידה האנליטית, נמצא שבשיטה זו ממד המבנה תלוי בשני הממדים האחרים במידה כזו שאין הצדקה להעריך אותו בנפרד, ולמעשה מוערכים רק שני ממדים – תוכן ולשון. לנוכח ממצאים אלו גובש מדריך נקידה חדש, בעל שני ממדים: ממד התוכן (הכולל גם מאפיין מבני אחד) וממד הלשון. מדריך זה, המוצג להלן, שימש את מפתחי הפרויקט במרבית ההערכות בשנתו האחרונה (1997), ולדבריהם הוכיח את עצמו כמוצלח ומתאים לצרכי הבחינה.

מדריך נקידה דו-ממדי

על המעריך לתת לכל חיבור שני ציונים על שני ממדים נפרדים של החיבור : תוכן ולשון.

בכל ממד יש לתת ציון על סולם בין 1 (הציון הנמוך ביותר) ל-6 (הציון הגבוה ביותר).

במקרים נדירים יימצאו תוצרי כתיבה קצרים בהרבה ממינימום השורות הנדרש, ובכל זאת מעולים. ברוב המקרים, הקיצור פוגע באיכותם של שני הממדים, והנקידה תיעשה בהתאם לכך. במקרים של כתיבה בהיקף מינימלי, שבהם המעריך אינו בטוח מספיק במהימנות הערכתו, לא יינתן ציון.

תוכן -

1.....2.....3.....4.....5.....6

* האם ברור לקורא מה רצה הכותב לומר (ללא צורך במידה רבה של פרשנות מצד הקורא)?

* האם הכתיבה מעידה על הבנה של הקטע או הטבלה שעליהם התבססה משימת הכתיבה?

* האם הכותב ענה על משימת הכתיבה? האם התייחס לכל מרכיבי המשימה?

* האם תוכן הכתיבה רלוונטי, מפורט ומקיף דיו?

* האם ניכרת הבחנה טובה בין עיקר לטפל?

* האם אין חזרות והארכות מיותרות?

* האם הכותב עקיב בדבריו (כלומר, דבריו אינם מכילים סתירות פנימיות לא-מטופלות)?

* האם דבריו של הכותב מבוססים ומשכנעים (במשימות שבהן מתבקשת נקיטת עמדה)?

* במקרים שבהם הכותב משתמש בדוגמאות – האם הן ממלאות את תפקידן בטקסט (שכנוע, הבהרה וכדומה)?

* האם בקריאת הטקסט כיחידה אחת אפשר להבחין בפיתוח מאורגן ורציף של התוכן?

לשון -

1.....2.....3.....4.....5.....6

* האם הטקסט רציף ומקושר (מלוכד)? האם יש שימוש במילות קישור מתאימות בין משפטים ובין פסקאות?

* האם אוצר המילים וסגנון הכתיבה מתאימים לשפה כתובה ולסגנון כתיבה עיוני?

* האם המשפטים תקינים מבחינה תחבירית (מבנה המשפט, סדר מילים, מילות יחס וכו')?

* האם נעשה שימוש סמנטי מדויק במילים?

* האם נשמרים כללי הלשון ואין שגיאות דקדוקיות חוזרות?

* האם הניסוח מדויק ובהיר?

בנוסף להיות ההערכה אנליטית, יש כמה מאפיינים אחרים שיש לזכור בעת ההתייחסות למחווון זה:

אופי המטלות

המטלות הנכללות במבחן הכתיבה שפותח במאל"ו מבוססות על גריינים ארוכים יחסית, שמופיעים באחד מכמה פורמטים: קטעי קריאה בודדים, זוגות של קטעי קריאה או תרשימים המוסרים מידע. שלא כמו במטלות במבחנים בארצות הברית, במטלות אלו יש ניסיון למזער את ההסתמכות על הידע המוקדם של הנבחן, באמצעות מתן מידע עשיר יחסית ודרישה להתייחס בחיבורים, לפחות חלקית, למידע זה.

שיטת הנקידה

כל מטלה מוערכת הערכה אנליטית בשני ממדים, תוכן ולשון – כל אחת מההערכות בסולם 1-6, במדרגות של נקודה. שני מעריכים בלתי תלויים נותנים למטלה ציון בכל אחד מהממדים, וסכום הציונים שנתנו הוא ציון הגלם של הנבחן. כלומר, טווח הציונים הוא 4-24.

אופן ההנחיה למעריך

כאמור, המדריך שפותח במאל"ו מציג את שני הממדים שיש להעריך את החיבורים ואת מרכיביהם על פיהם, אך אינו מפרט פירוט איכותי של כל אחת מרמות הציון בכל אחד מהממדים.

הצגה גרפית

שלא כמדריכי הנקידה של ה-SAT ושל ה-GRE, המאורגנים בטבלאות (בגלל אופן ההנחיה, המפרט את רמת המדדים האופיינית לכל ציון), כאן מופיעות שתי רשימות של שאלות.

תהליך פיתוח המחווון

הוליסטי או אנליטי?

להכרעה בין שתי שיטות ההערכה יש השפעות רבות, הנוגעות להיבטים שונים של המטלה בפרט ושל הבחינה בכלל. לאחר שקילת היתרונות והחסרונות של כל שיטה, הוחלט לבחור בשיטה האנליטית. החלטה זו קיבלה חיזוק מבדיקה שערך צוות הפיתוח לפני הניסוי הראשון. הצוות בדק את הערכותיהם של 76 חיבורים באותו נושא. כל אחד מהחיבורים הוערך שלוש הערכות בעזרת מחווון הוליסטי ושלוש הערכות בעזרת מחווון אנליטי (בשני ממדים, תוכן ולשון), וכל ההערכות ניתנו בסולם 1-3. המעריכים היו אנשי צוות הפיתוח, וכל מעריך התנסה בשתי השיטות. ממוצע המתאמים בין שלשות המעריכים ההוליסטיים, 0.462, היה נמוך מממוצע המתאמים בין שלשות המעריכים האנליטיים – 0.624. מגמה דומה נרשמה גם בהעברות הניסיוניות של מטלות מבחן הכתיבה שפותח במאל"ו בשנות התשעים. ארבע מטלות כתיבה הועברו אז העברה ניסיונית בשתי שיטות הערכה – הוליסטית ואנליטית. ממוצע המתאמים בין המעריכים בשיטה ההוליסטית היה 0.52, ואילו ממוצע המתאמים בין המעריכים בשיטה

האנליטית היה 0.60. בקרב המעריכים האנליטיים, המתאם בין הערכת ממד התוכן להערכת ממד הלשון אצל אותו מעריך היה נמוך באופן מפתיע – 0.292.

שימוש במחווון האנליטי שפותח במאל"ו או פיתוח מחווון חדש?

צוות הפיתוח החליט לא להתבסס רק על המחווון שפותח במאל"ו בשלהי שנות התשעים, אלא לפתח מחווון אנליטי שישלב את יתרונותיהם של המחווונים משתי הגישות שהוצגו לעיל. כאמור, יתרונותיה של הגישה האנליטית שבמדריך הנקידה של מאל"ו ברורים, ומנגד, לשיטת ההנחיה למעריך במדריכי הנקידה האמריקניים – שבהם פירוט איכותי של כל אחת מרמות הציון בכל אחד מהממדים – יתרון בולט בהפחתה של רעש פוטנציאלי ובהגדלת המהימנות. הצורך במחווון חדש נובע גם מהעובדה שהגריינים ומטלות הכתיבה שפותחו בידי צוות הפיתוח ושהוצגו לעיל שונים הן מהמטלות הקצרות המשמשות במבחנים האמריקניים והן מהמטלות הארוכות ששימשו במבחן שפותח בעבר במאל"ו.

ניתוח תוכן ואופרציונליזציה של יכולת הכתיבה

פיתוח המחווון החדש הצריך בחירה וניסוח של קריטריונים שישמשו להערכה בכל אחד מממדי ההערכה. לשם כך נערך ניתוח של יכולת הכתיבה, מתוך שאיפה להגיע להגדרה ממצה של מרכיביה שיכולים להימדד בהקשר הנתון ובמגבלות הבחינה (מגבלת הזמן, שנגזרת ממנה גם מגבלה בהיקף החומר שאפשר להציג לנבחנים, וכן המגבלה של מבחן נייר ועיפרון). מרכיבים אלו נאספו ממחקרים בנושא, וכבסיס ראשוני להתייחסות שימשה רשימת המרכיבים שבמחווון שפותח במאל"ו בעבר.

מחקרים בנושא מציגים בדרך כלל עד חמישה ממדים שונים להערכת יכולת הכתיבה: תוכן, מבנה, דקדוק ותחביר, סגנון ואוצר מילים. את שלושת האחרונים אפשר להכיל יחד בממד "לשון". המחווון שנבנה במאל"ו בשלהי שנות התשעים שיקף ניסיון להקיף את שלושת ממדי הכתיבה – תוכן, מבנה ולשון. ואולם, מפתחי המבחן מצאו כי יש מתאם גבוה (0.7-0.9) בין ממדי התוכן והמבנה, מתאם בינוני (0.6-0.8) בין ממדי המבנה והלשון, ומתאם נמוך יותר (0.4-0.7) בין ממדי התוכן והלשון. כלומר, ממד המבנה לא היה מובחן מספיק מממד התוכן, ואף שהמחווון שפותח היה תלת-ממדי, למעשה הוערכו בעזרתו שני ממדים בלבד – תוכן ולשון. נוכח ממצאים אלה, וגם משיקולי תקציב, הוחלט שהמחווון החדש שיפותח יכלול שני ממדים – תוכן ולשון – וכל אחד מהם יכלול גם מאפיינים מבניים רלוונטיים לו (בממד התוכן – מיקוד ולכידות, בממד הלשון – שימוש באמצעי ארגון).

לאחר שגובשה רשימה ראשונית של מרכיבי יכולת הכתיבה הרלוונטיים, החל תהליך ממושך של התאמתם לאופי המטלות, למטרתן (הערכת יכולת כתיבה אקדמית) ולמגבלות ההערכה, ולבסוף נבחרו ונוסחו המרכיבים הנכללים בכל ממד הערכה. לתהליך זה תרם רבות צוות הליווי מהיחידה להוראת כתיבה מדעית במכינה הקדם-אקדמית ובפקולטה למדעי הרוח באוניברסיטה העברית.

המרכיבים הנכללים בממד התוכן הם :

- קיומו של רעיון או ציר רעיוני ברור, מבוסס ורלוונטי למטלה
- מידת הקשר בין הרעיון לפיתוחו
- מיקוד ולכידות של התוכן
- ביטוי לחשיבה ביקורתית

המרכיבים הנכללים בממד הלשון הם :

- בהירות
- הלימה לכתיבה עיונית
- דיוק סמנטי
- תקינות לשונית
- שימוש באמצעי ארגון לשוניים
- גיוון תחבירי

השלב הבא היה ניסוח אופרטיבי – תיאור ממצה של ביטוייה של התכונה הנמדדת – בכל אחת משש רמות הציון. פותח מחוון המורכב משתי טבלאות בלתי תלויות, ובכל אחת מהן פירוט בנוגע למאפיינים האיכותיים של כל אחת מרמות הציון. בכל רמת ציון מוקדש התא העליון (המודגש) לאפיון כללי של חיבור ברמה זו. לפני שתי הטבלאות מופיעות הנחיות כלליות לשימוש במחוון.

שיטת הנקידה

שיטת הנקידה שנבחרה היא השיטה ששימשה במבחן שפותח במאל"ו בעבר, ושסולם הציונים בה זהה למעשה לזה שבבחינות האמריקניות. במהלך הפיתוח נבחנה גם האפשרות להשתמש במודל אנליטי דו-ממדי מצומצם. במודל זה כל מטלה מוערכת הערכה אנליטית בשני ממדים – תוכן ולשון – על ידי שני מעריכים, וכל אחת מההערכות היא בסולם 1-3 (חלש, בינוני, טוב). הציון במטלה הוא סכום שתי ההערכות, ומטלות טובות במיוחד זוכות בנקודה נוספת בהערכה הסופית של כל מעריך. טווח הציונים הוא אפוא 4-14. אפשרות זו הועלתה בשל החשש מאי-יכולת להבחין בין ציונים ברמות הביניים, אולם הוחלט לוותר עליה, מפני שהסולם שהיא מספקת גס מדי.

הוחלט שתגובה בלתי סבירה למטלה תקבל ציון 0 בשני הממדים. שלושה סוגי חיבורים הוגדרו תגובה בלתי סבירה : חיבור שאינו מתייחס למטלה (להבדיל מחיבור המפגין יכולת נמוכה ביותר, אך יש בו בכל זאת התייחסות למטלה), חיבור הכתוב בשפה שונה מהנדרש וחיבור שאורכו פחות מ-10 שורות.

הוחלט גם שלחיבורים קצרים מהנדרש (שאורכם 10-23 שורות) תופחת נקודה אחת מהציון בממד הלשון. ההנחה היא שבחיבורים קצרים כל כך התוכן יהיה דל ממילא, שכן לפחות עד אורך מסוים – הגדול במידה ניכרת מאורך המינימום הנדרש – קיים יחס ישר בין אורך החיבור ובין רמתו ומורכבותו התוכנית, ולכן אין סיבה "לקנוס" חיבור כזה ולהוריד לו ניקוד נוסף בממד התוכן. בממד הלשון, לעומת זאת, סביר שיופיעו בחיבורים קצרים פחות טעויות לשוניות משהיו

מופיעות אצל אותו כותב אילו היה עומד בהיקף הכתיבה הנדרש, וההפחתה נועדה למנוע הערכת יתר של ממד זה שעלולה להיגרם בגלל הקיצור.

להלן המחווון בצורתו הסופית:

המחווון – מדריך הערכה דו-ממדי לחיבורים

עליכם להעריך את החיבורים בשני ממדים: ממד התוכן וממד הלשון.

יש להעריך כל ממד באופן שאינו תלוי בהערכת הממד האחר.

בהערכת ממד התוכן יש להתייחס לקריטריונים האלה: קיומו של רעיון או ציר רעיוני ברור, מבוסס ורלוונטי למטלה; חוזק הקשר בין הרעיון לפיתוחו; המיקוד והלכידות של התוכן; ביטוי לחשיבה ביקורתית.

בהערכת ממד הלשון יש להתייחס לקריטריונים האלה: בהירות; הלימה לכתיבה עיונית; דיוק סמנטי; תקינות הלשון; שימוש באמצעי ארגון לשוניים; גיוון תחבירי.

יש להעריך כל ממד בסולם ציונים 1-6. בכל ממד עליכם לבחור מהטבלה שלהלן בציון המתאר בדרך הטובה ביותר את החיבור שקראתם.

הערות:

* במקרים רבים שום תיאור לא יתאים באופן מושלם לחיבור שתעריכו. במקרים כאלה יש לבחור בתיאור הקרוב ביותר ברוחו לחיבור.

* במקרים שבהם החיבורים אינם תואמים את ההוראות שניתנו לנבחנים יש לפעול כמפורט להלן:

* חיבורים שכתובים בשפה שאינה שפת הבחינה יש לשלוח בחזרה לאחראי האדמיניסטרטיבי לבדיקה.

* בחיבורים שאורכם 10-23 שורות יש לבחור באפשרות "כן" כתשובה לשאלה "אורך החיבור 10-23 שורות?". לאחר מכן יש להעריכם.

* בחיבורים החורגים מהשורות הנתונות בגיליון הכתיבה, השורות החורגות לא ייקראו ולא יובאו בחשבון בעת ההערכה.

* בחיבורים שתוכנם אינו מתאים כלל לנושא המטלה, או בחיבורים שאורכם 0-9 שורות, יש לבחור באפשרות "לא" כתשובה לשאלה "כתיבה לנושא?". בחיבורים אלה יש לסמן 1 הן בממד התוכן הן בממד הלשון.

תוכן

1 - חלש מאוד	2 – חלש	3 - בינוני-נמוך	4 - בינוני-גבוה	5 - טוב	6 - מעולה
אין רעיון מרכזי או ציר רעיוני ברור, או יש זיקה מועטה של תוכן החיבור אל המטלה	רעיון מרכזי או ציר רעיוני לא מבוסס ולא מפותח, או זיקה לא מספקת של תוכן החיבור אל המטלה	רעיון מרכזי או ציר רעיוני שהביסוס והפיתוח שלו מוגבלים, או זיקה לא מלאה של תוכן החיבור אל המטלה	רעיון מרכזי או ציר רעיוני מבוסס ומפותח בצורה טובה למדי	רעיון מרכזי או ציר רעיוני מבוסס ומפותח כהלכה	רעיון מרכזי או ציר רעיוני מפותח ומבוסס באופן יוצא דופן
התוכן דל	פיתוח דל של הרעיון, או קשר מועט בין הרעיון לפיתוחו	קשר לא חזק דיו בין הרעיון לפיתוחו (הסברים לא מספקים, טיעונים לא משכנעים או דוגמאות לא מתאימות)	פיתוח רלוונטי לרעיון (הסברים מספקים ברובם, טיעונים משכנעים למדי, דוגמאות מתאימות)	פיתוח רלוונטי ועשיר של הרעיון (הסברים מלאים, טיעונים משכנעים, דוגמאות מתאימות)	פיתוח רלוונטי, מעמיק ועשיר של הרעיון (הסברים מלאים, טיעונים משכנעים ומגוונים, דוגמאות מתאימות)
היעדר מיקוד ולכידות (קפיצות מעניין לעניין ללא כל שמירה על רצף רעיוני, חזרות מיותרות)	בעיות קשות במיקוד ובלכידות (שמירה מועטה ביותר על רצף רעיוני, חזרות מיותרות)	המיקוד והלכידות מוגבלים למדי (שמירה מועטה על רצף רעיוני)	התוכן ממוקד ולכיד למדי (שמירה על רצף רעיוני לאורך רוב החיבור, הצגה מדויקת בדרך כלל של הקשרים בין הרעיונות שבחיבור)	התוכן ממוקד ולכיד (שמירה על רצף רעיוני לאורך כל החיבור, הצגה מדויקת של הקשרים בין הרעיונות שבחיבור)	התוכן ממוקד ולכיד ביותר (שמירה על רצף רעיוני לאורך כל החיבור, הצגה מדויקת וממצה של הקשרים בין הרעיונות שבחיבור)
אין ביטוי לחשיבה ביקורתית *	ביטוי מועט ביותר לחשיבה ביקורתית *	ביטוי מועט למדי לחשיבה ביקורתית *	ביטוי לחשיבה ביקורתית *	ביטוי מרשים לחשיבה ביקורתית *	ביטוי מרשים מאוד לחשיבה ביקורתית *

* הגדרה מדויקת של הסוגיה, הבחנה בין דעה לעובדה, בחינה של סוגיות מכמה נקודות מבט, התמודדות עם עמדות מנוגדות

לשון

1 - חלש מאוד	2 – חלש	3 - בינוני-נמוך	4 - בינוני-גבוה	5 - טוב	6 - מעולה
הלשון עילגת ודלה במיוחד	הלשון דלה ואינה בהירה	הלשון אינה בהירה דיה	הלשון בהירה למדי	הלשון בהירה	הלשון בהירה, קולחת ועשירה
ניסוח שאינו הולם כלל כתיבה עיונית	ניסוח שברובו אינו הולם כתיבה עיונית	ניסוח שבמקומות לא מועטים אינו הולם כתיבה עיונית	ניסוח ההולם ברובו כתיבה עיונית	ניסוח ענייני ההולם כתיבה עיונית	ניסוח ענייני ההולם כתיבה עיונית
דוגמאות מרובות לשימוש סמנטי לקוי במילים	דוגמאות לא מעטות לשימוש סמנטי לקוי במילים	דוגמאות מעטות לשימוש סמנטי לא מדויק במילים	שימוש סמנטי מדויק ברובו במילים	שימוש סמנטי מדויק במילים	שימוש סמנטי מדויק במילים
שגיאות דקדוק ותחביר רבות	דוגמאות לא מעטות לשגיאות דקדוק או תחביר	דוגמאות מעטות לשגיאות דקדוק או תחביר	דוגמאות בודדות בלבד לשגיאות דקדוק או תחביר	דקדוק ותחביר תקינים	דקדוק ותחביר תקינים
שימוש במבנים תחביריים בסיסיים ביותר	שימוש במבנים תחביריים לא מורכבים ולא מגוונים	שימוש במבנים תחביריים לא מורכבים ולא מגוונים בדרך כלל	דוגמאות ספורות לשימוש הולם במבנים תחביריים מורכבים למדי	דוגמאות רבות לשימוש הולם במבנים תחביריים מורכבים	שימוש נרחב במבנים תחביריים מורכבים ומגוונים, התורמים למשמעות הכתוב
היעדר אמצעים לשוניים המארגנים את הכתוב, או שימוש לקוי באמצעים כאלה *	דוגמאות מעטות ביותר לשימוש באמצעים לשוניים המארגנים את הכתוב, או שימוש לקוי ברובו באמצעים כאלה *	דוגמאות מעטות לשימוש באמצעים לשוניים המארגנים את הכתוב, או דוגמאות מעטות לשימוש לקוי באמצעים כאלה *	דוגמאות לא מעטות לאמצעים לשוניים המארגנים את הכתוב *	דוגמאות רבות לשימוש באמצעים לשוניים המארגנים את הכתוב *	שימוש נרחב ומגוון באמצעים לשוניים המארגנים את הכתוב*

* מילות קישור, משפטי מעבר, חלוקה לפסקאות

הערכות החיבורים – ניסוי שני

שלא כמו בניסוי הראשון, שהמעריכים השתמשו בו בשני מחוונים – הוליסטי (מחוון ה-SAT) ואנליטי (המחוון שפותח במאל"ו בשנות התשעים) – בניסוי השני השתמשו המעריכים במחוון האנליטי הדו-ממדי החדש. מחוון זה היה גלגול מוקדם של המחוון שהוצג לעיל, שעודכן מאז – בין השאר בעקבות ממצאי הניסוי ומשוב המעריכים. את חיבורי הניסוי העריכו 12 מעריכים, שחולקו לשתי קבוצות. כל קבוצה העריכה 4 מטלות, וכל מעריך – כ-150 חיבורים.

טבלה מספר 6 – המתאמים בין ממדים ובין מעריכים (בכל המטלות) בניסוי השני

	מעריך 1 - תוכן	מעריך 1 - לשון	מעריך 2 - תוכן	מעריך 2 - לשון
מעריך 1 - תוכן	1.00	0.71	0.50	0.42
מעריך 1 - לשון		1.00	0.43	0.51
מעריך 2 - תוכן			1.00	0.65
מעריך 2 - לשון				1.00

טבלה מספר 7 – המתאמים בין ממדים ובין מעריכים במטלות המקוריות בעברית (6 מטלות)

	מעריך 1 - תוכן	מעריך 1 - לשון	מעריך 2 - תוכן	מעריך 2 - לשון
מעריך 1 - תוכן	1.00	0.71	0.53	0.46
מעריך 1 - לשון		1.00	0.46	0.54
מעריך 2 - תוכן			1.00	0.64
מעריך 2 - לשון				1.00

בנוסף לנתונים שבטבלאות, המתאם הכללי בין הערכות התוכן ובין הערכות הלשון של אותו מעריך היה 0.73. כלומר המתאם בין הערכות הממדים השונים אצל אותו מעריך היה גבוה מהמתאם בין הערכות של מעריכים שונים לאותו ממד.

מגמה זו קיבלה ביטוי קיצוני יותר במטלות המתורגמות מערבית, שבהן מתאמי הערכות התוכן בין שני המעריכים היו נמוכים במידה ניכרת לעומת המטלות המקוריות.

טבלה מספר 8 – המתאמים בין ממדים ומעריכים במטלות המתורגמות מערבית (2 מטלות)

	מעריך 1 - תוכן	מעריך 1 - לשון	מעריך 2 - תוכן	מעריך 2 - לשון
מעריך 1 - תוכן	1.00	0.67	0.36	0.31
מעריך 1 - לשון		1.00	0.31	0.43
מעריך 2 - תוכן			1.00	0.68
מעריך 2 - לשון				1.00

נראה שנתונים אלה מעידים כי על אף ההנחיה להערכה בלתי תלויה של כל אחד מממדי החיבורים, יש בהערכה יסוד הוליסטי דומיננטי. כזכור, בבדיקה שנועדה להשוות בראשית תהליך הפיתוח בין הגישה ההוליסטית לאנליטית, ממוצע המתאמים בין המעריכים האנליטיים היה 0.624, והמתאם בין הערכת ממד התוכן להערכת ממד הלשון אצל אותו מעריך היה 0.292.

בשנים הקרובות תיערך השוואה בין הערכה אנליטית של חיבורי אמת לבין הערכה הוליסטית שלהם. אם לא תימצא עדיפות להערכה האנליטית, יישקל מעבר להערכה הוליסטית, משיקולי עלות ההערכה.

יש לציין כי בחינת המתאמים בין הזוגות השונים של המעריכים מגלה שטווח המתאמים גדול – בין 0.28 ל-0.86. שלא כמעריכי המטלות התפעוליות, מעריכי מטלות הניסוי לא עברו סלקציה המבוססת על ביצועיהם, ויש להניח שהליך הסינון וההכשרה, וכן הניסיון המצטבר של מעריכים ותיקים, יביאו לעלייה במתאמים.

הכשרת המעריכים

להליך ההכשרה שעוברים מעריכים השפעה ניכרת על מהימנות הערכותיהם: מהימנות ההערכה עולה ככל שההכשרה מעמיקה ויסודית יותר (Freedman, 1981). כדי להבטיח הערכה מהימנה ותקפה ככל האפשר הוחלט להשקיע משאבים רבים בפיתוח הליך ההכשרה של המעריכים.

הליך הקבלה וההכשרה

המועמדים עוברים סינון ראשוני, לפי המלצות וקורות חיים. נתוני הסף הם אלה:

- השכלה: תואר אקדמי ראשון לפחות.
- ניסיון מקצועי: ניסיון בהוראת שפת אם או בעריכת לשון או בכתיבה עיתונאית/ספרותית או בתרגום חומר כתוב או בהגהת בחינות.

בוגרי החוגים ללשון עברית/ערבית או החוגים לבלשנות פטורים מן הדרישה לניסיון מקצועי.

המועמדים שעברו את הסינון הראשוני מקבלים בדוא"ל חומר רקע ובו מידע כללי על דרך העבודה (דגשים בהערכה, הערכות של חיבורים נבחרים, הנחיות לשימוש במחווך וכולי). חומר זה, יחד עם החומר שניתן להם במא"ל, משמש בסיס לשלב הבא: התנסות מעשית הכוללת כתיבת חיבור והערכה של כמה חיבורים בנושא נתון. התנסות זו היא למעשה מבחן קבלה, ועל בסיס תוצאותיה מתקיים סינון שני. המתקבלים עוברים סדנת הכשרה לפני כל מחזור בחינה.

סינון נוסף ייערך על סמך ביצועי המעריכים בהערכות התפעוליות. לאחר כל מועד ייבחנו המתאמים של כל מעריך עם יתר המעריכים, ונתונים נוספים הקשורים בהערכתו (למשל מספר ההערכות השלישיות שנגרמו "באשמתו"). המעריכים שביצועיהם לא יהיו משביעי רצון לא יעריכו שוב.

סדנת הכשרה למעריכים

פותח פורמט סדנה מקיפה להכשרת מעריכים.

בשלב הראשון פותחה סדנה פרונטלית על בסיס הסדנה להכשרת המעריכים בבחינת יע"ל. החלק הראשון של הסדנה מוקדש להכשרה כללית, ומוצגים בו בפרוטרוט אופי מטלת הכתיבה, הגישה הנכונה להערכה דו-ממדית, אופן העבודה עם המחווך והיבטים טכניים שונים של העבודה. עיקר הסדנה מיוחד למטלה הספציפית שהמשתתפים בסדנה יעריכו את החיבורים שנכתבו בתגובה לה. חלק זה כולל בעיקר עבודה מעשית: הערכה של חיבורים וניתוח שלהם. החיבורים המוצגים מוערכים מראש על ידי כמה חברים מצוות הפיתוח, ומטרת הסדנה היא לכייל את הערכותיהם של המעריכים המתלמדים לציונים שניתנו לחיבורים על ידי חברי הצוות. בכל סדנה משתתפים כ-10-15 מעריכים.

להלן דוגמה לניתוח גרפי של חיבור המוצג בסדנה:

חיבור לדוגמה - לשון

הפתיחה מנותקת מהחיבור.

לשון לא בהירה דיה - ניסוחים שבהם סדר מילים משבש שפוגע במשמעות.

(1) לשון עילגות. ניסוח שאינו הולם כתיבה עיונית.

(5) לשון עילגות. ניסוח שאינו הולם כתיבה עיונית.

(10) יש שימוש באמצעים לשוניים לארגון הכתוב.

לשון החיבור אינה בהירה דיה, במקומות לא מעטים הניסוח אינו הולם כתיבה עיונית. יש דוגמאות לשימוש באמצעים לשוניים המארגנים את הכתוב.

חיבור לדוגמה - תוכן

הטיעון המרכזי לא מבוסס. אין הבחנה בין דעה לעובדה.

קיים רעיון מרכזי, אך הביסוס והפיתוח שלו מוגבלים.

עוד קביעה ללא ביסוס. (5)

טיעון לא חזק מספיק (אין סיבה).

קביעה לא מבוססת, ולא מובחנת מספיק מהטיעון הקודם. (10)

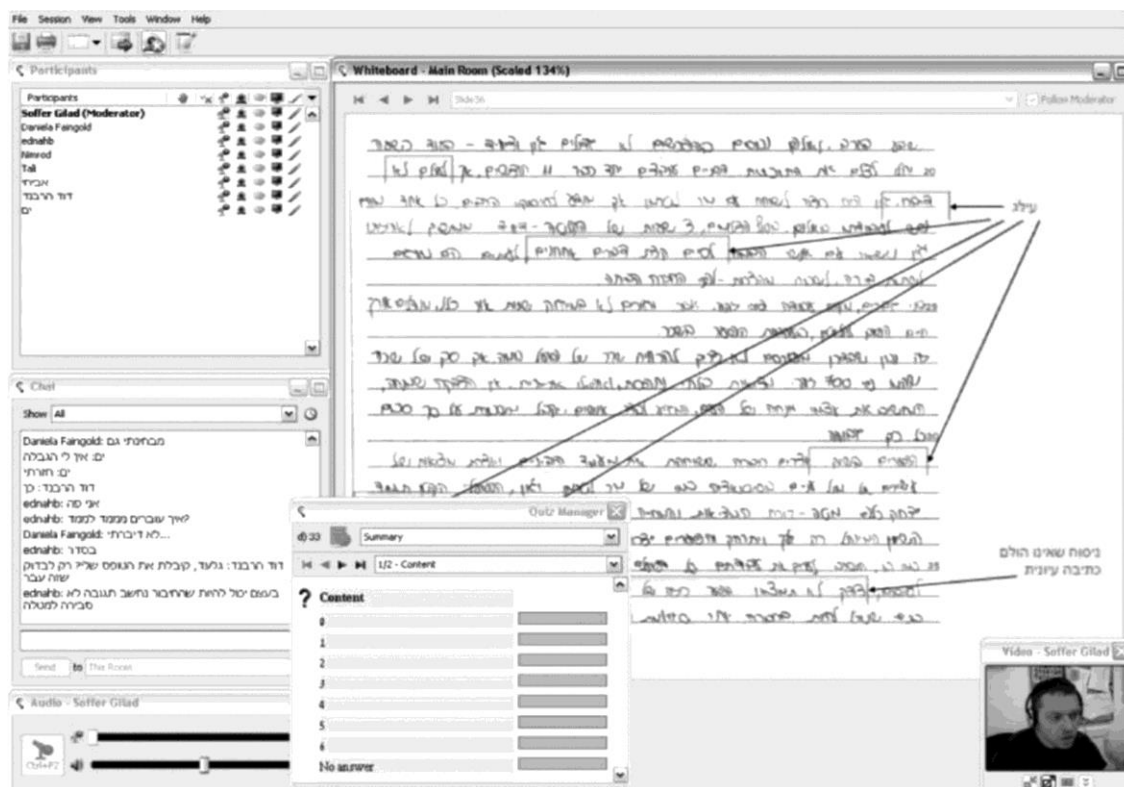
רצף רעיוני מוגבל (הנביעה אינה מתחייבת).

פרשנות צרה של ערכים.

סדנה מתוקשבת

על בסיס הסדנה הפרונטלית נבנתה סדנת הכשרת מעריכים מתוקשבת. הסדנה מועברת באמצעות תוכנת e-learning ששמה blackboard. הסדנה המתוקשבת דומה לסדנה הפרונטלית בתכניה, אך התקשורת בין המנחה ובין המשתתפים שונה – התקשורת ספונטנית פחות, וחד-כיוונית ומובנית יותר.

המטרה היא שבעתיד מעריכים ותיקים יוכלו לעבור סדנאות מתוקשבות, וכך יוכלו לעבור את הסדנה בביתם.



צילום מסך מתוך סדנה מתוקשבת

השוואה בין הסדנאות הפרונטליות והמתוקשבות
טבלה מספר 9 – השוואה של המתאמים בין המעריכים, על פי סוג הסדנה, בניסוי השני

0.53	סדנאות פרונטליות (6 סדנאות, n=708)
0.63	סדנאות מתוקשבות (2 סדנאות, n=221)
0.58	סדנאות פרונטליות – מטלות מקוריות בלבד (4 סדנאות, n=495)

המתאמים בין המעריכים שעברו סדנאות מתוקשבות היו גבוהים יותר, אך סדנאות אלו נערכו אחרי שכל מעריך עבר כבר שלוש סדנאות פרונטליות, ולכן ייתכן שמקור הפער במתאמים בין הסדנאות הפרונטליות למתוקשבות נובע מן הניסיון המצטבר של המעריכים.

טבלה מספר 10 – השוואה של המתאמים בין הציון במטלת הכתיבה ובין ציוני האמת בבחינה בתחום המילולי ובבחינה כולה, בניסוי השני

	תוכן	לשון	כללי	
סדנאות פרונטליות (6 סדנאות, n=674)	VE	0.44	0.52	0.52
	TOT	0.42	0.48	0.48
סדנאות מתוקשבות (2 סדנאות, n=211)	VE	0.40	0.52	0.49
	TOT	0.44	0.52	0.52
סדנאות פרונטליות – מטלות מקוריות בלבד (4 סדנאות, n=468)	VE	0.49	0.56	0.56
	TOT	0.47	0.52	0.53

נמצא יתרון קל לסדנאות הפרונטליות (כשמתאמים למטלות המקוריות בלבד). יש להתייחס לנתונים הללו בזהירות, שכן המתאמים מושפעים מהגריינים הספציפיים, מהמעריכים הספציפיים, ממעבירי הסדנאות הספציפיים וכולי.

אופן הערכת החיבורים בפועל

בשני הניסויים נעשו הערכות החיבורים במאל"ו, והמעריכים ראו צילומים של החיבורים, ללא פרטים מזהים. ההערכה התפעולית, לעומת זאת, מתבצעת באמצעות מחשב: לאחר העברת הבחינה הפסיכומטרית, כל החיבורים נסרקים ומועלים למערכת הערכת חיבורים ממוחשבת, מערכת סופרסטאר.

משהשתתף מעריך בסדנה העוסקת בגריין מסוים, נפתחת בפניו האפשרות להתחבר למערכת סופרסטאר ממחשבו האישי בבית, דרך רשת האינטרנט. המערכת מפנה אליו חיבור אקראי. על

המעריך למלא 4 שדות : הערכת תוכן, הערכת לשון, קביעה אם החיבור קצר (10-23 שורות) וקביעה אם החיבור פסול .

המעריכים יכולים להפנות שאלות טכניות לאחראי אדמיניסטרטיבי (למשל בנוגע לחיבור שלא נסרק באיכות טובה דיה, או בנוגע לכל תקלה שהיא במערכת), ושאלות הנוגעות להערכה עצמה – לאחראי תוכן.

חלק ד – סוגיות נוספות בפיתוח מטלת הכתיבה במכפ"ל

היקף הכתיבה במטלת הכתיבה

במהלך פיתוח המטלות עלתה השאלה אם יש להגביל את היקף הכתיבה בחיבור בחסם עליון ו/או בחסם תחתון. הצבת חסם תחתון מבטיחה שהדגימה שיכולות הכתיבה של הנבחנים מושוות בהתבסס עליה תהיה רחבת היקף דיה, ומצמצמת את הפגיעה במהימנות שנובעת מפרשנות שונה של דרישות המטלה. הצבת חסם עליון מונעת שונות גדולה מדי בהיקף החיבורים, וכן מקלה על העיצוב והסריקה של תשובוני הבחינה.

הוחלט להגדיר את היקף החיבור הנדרש על פי מספר השורות הנכתבות, ולא על פי מספר המילים, כדי למנוע מהנבחנים (ומהמעריכים) את האי-נוחות הכרוכה בספירה או בחישוב של מספר המילים בחיבור. בדיקות שנערכו בשני הניסויים העלו כי דרישת הסף המינימלית המאפשרת הערכה הוגנת של יכולת הכתיבה היא 25 שורות. עם זאת, מחקרים רבים מראים יחס ישר בין אורכי חיבורים ובין הערכותיהם, ולכן בחוברת ההסבר מובהר לנבחנים שרוב החיבורים הזוכים להערכות גבוהות הם ארוכים. הוחלט לא לציין מפורשות היקף כתיבה מקסימלי, אך להבהיר כי אין לחרוג מן השורות המוקצות לכתיבה וכי לא יינתן גיליון כתיבה נוסף. לפיכך, בפועל, טווח הכתיבה הלגיטימי הוא 25-50 שורות.

כפי שצוין בהתייחסות לנקידת החיבורים (חלק ג לעיל), חיבורים קצרים מ-25 שורות לא ייקנסו בממד התוכן, מתוך הנחה שחיבורים כאלה יהיו דלים בתוכנם ממילא. עם זאת, בשל טבעה של שיטת הנקידה בממד הלשון (בעיקר הפחתת נקודות, ולא צבירת נקודות, כבממד התוכן), קוצרו של החיבור יכול לפעול דווקא לזכותו בממד זה, ולפיכך הוחלט שחיבורים הקצרים מ-25 שורות ייקנסו, ומציונם בממד הלשון תופחת נקודה. חיבורים שאורכם 9 שורות ומטה ייפסלו.

ביסוס החסם התחתון לכתיבה על מספר שורות מאפשר רק אומדן של היקף הכתיבה המינימלי שיתקבל, שכן מספר המילים בכל שורה נגזר מגודל כתב היד של הנבחן. בדיקת גודל כתב היד של נבחנים אקראיים שנבחנו בניסוי השני (50 חיבורים בעברית ו-50 בערבית) העלתה כי אכן קיימת בו – ומכאן, גם במספר המילים שבחיבורים – שונות רבה. טווח הכתיבה שהוגדר בניסוי היה 25 עד 50 שורות. בחיבורים בעברית היה מספר המילים המקסימלי בחיבור 409, והמינימלי – 126 (בערבית – 527 ו-121). כך נוצר מצב שבו חיבורים העומדים בהיקף הכתיבה המינימלי – 25 שורות – עלולים להכיל פחות מילים מחיבורים שאינם עומדים בדרישת הסף. ואכן, מספר המילים המינימלי בחיבור שעמד בדרישת הסף היה 176, ומספר המילים המקסימלי בחיבור שלא עמד בדרישת הסף היה 222 (בערבית 178 ו-213 בהתאמה). עם זאת, נוכח החלופות הקיימות (ספירת מילים או הימנעות מהגבלת היקף הכתיבה) נראה כי הגדרת היקף החיבור הנדרש על פי מספר השורות הנכתבות הוא הפתרון הטוב ביותר בנמצא כרגע כל עוד הבחינה אינה ממוחשבת.

בחירה

במטלות כתיבה נהוגים שני סוגי בחירה עיקריים: בחירה בתוך מטלה המתייחסת לתחום נתון באופן כללי (למשל: "ציין אישיות שאתה מעריך ונמק מדוע אתה מעריך אותה"), ובחירה בין מטלות שונות לחלוטין. שימוש בסוג הבחירה הראשון אינו רצוי, שכן הוא פוגע בסטנדרטיות של הבחינה וביכולת להשוות בין חיבורים שונים, שכן תוכני החיבורים והמבנה שלהם מושפעים במידה ניכרת מבחירת הנושא ומהידע הפרטי שמשמש בבחירה ובכתיבה. לסוג הבחירה השני יש יתרון ברור על פני מצב של היעדר בחירה – צמצום השונות הלא-רלוונטית הנגרמת בשל ההבדלים במידת ההיכרות של הנבחנים עם עולם התוכן של הגריין – אך במחקרים לא נמצא קשר בין מתן בחירה ובין איכות הביצוע. ייתכן שההסבר לכך הוא שהבחירה כרוכה בהשקעת זמן בהערכת החלופות השונות, זמן שיש לו חשיבות מיוחדת נוכח קוצר הזמן המוקצה למטלת הכתיבה במכפ"ל. בנוסף, גם בחירה מהסוג השני מקשה על הערכה סטנדרטית, שכן היא יוצרת למעשה כמה נוסחים, שאינם בהכרח שקולים. במרבית הבחינות שנסקרו אין בחירה (יוצאת הדופן היא בחינת הבגרות בהבעה, שם יש בחירה בין שישה נושאים, אולם בבחינה זו מוקצה לנבחנים זמן רב בהרבה לעומת מטלות הכתיבה בבחינות האחרות). מהטעמים שהוזכרו הוחלט לא לאפשר בחירה בין נושאים שונים במטלת הכתיבה במכפ"ל.

הערכה שלישית

בהערכת פריטים פתוחים, ובכלל זה מטלת הכתיבה, השאיפה היא שההפרש בין הערכותיהם של מעריכים שונים יהיה נמוך ככל האפשר. עם זאת, מטבע הדברים לעתים אין הסכמה בין המעריכים על רמת החיבור. כאשר הפער גדול מקובל להעריך את החיבור הערכה נוספת.

איזה הפרש בין ההערכות מצדיק הפניית חיבור להערכה שלישית?

טבלאות מספר 11,12 – פערים בין הערכות המעריכים בניסוי השני
(מטלות מקוריות; הנתונים דומים גם במטלות המתורגמות)

הפרש בין הערכות המעריכים לאחר סכימת שני הממדים (סולם 2-12)	% מכלל החיבורים (מצטבר)
כללי - הפרש 0	22%
כללי - הפרש 1 או פחות	52%
כללי - הפרש 2 או פחות	75%
כללי - הפרש 3 או פחות	88%
כללי - הפרש 4 או פחות	95%
כללי - הפרש 5 או פחות	98%

הפרש בין הערכות המעריכים בכל אחד מהממדים (סולם 1-6)	% מכלל החיבורים (מצטבר)
תוכן - הפרש 0	34%
תוכן - הפרש 1 או פחות	78%
תוכן - הפרש 2 או פחות	93%
תוכן - הפרש 3 או פחות	98%
לשון - הפרש 0	35%
לשון - הפרש 1 או פחות	79%
לשון - הפרש 2 או פחות	95%
לשון - הפרש 3 או פחות	99%

הוחלט להעביר להערכה שלישית חיבורים שבהם פער של יותר מ-4 נקודות בציון הגלם הכללי, החלטה שמשמעותה הערכה שלישית של 5% מהחיבורים. ההחלטה תיבחן מחדש אם שיעור החיבורים שיגיעו להערכה שלישית ישתנה באופן ניכר. בכל מקרה הוחלט שבחיבורים שניתנו להם שלוש הערכות, תושמט ההערכה הרחוקה ביותר מהממוצע של שלושתן.

הפניית חיבורים למעריך בכיר – מקרים שבהם אחת ההערכות קובעת שהחיבור פסול
 כאמור, חיבור פסול הוא חיבור שאינו תגובה סבירה למטלה (אינו מתייחס למטלה, נכתב בשפה שונה מהנדרש או קצר מ-10 שורות). בניסוי השני נבדקו 929 חיבורים, וב-37 מהם ניתנה לפחות הערכה אחת של "חיבור פסול". מתוך 37 החיבורים האלה, ב-34 הייתה הסכמה של שני המעריכים בנוגע להערכה, ורק ב-3 חיבורים (9%) לא הייתה הסכמה כזאת. הוחלט כי במקרים כאלה יועברו החיבורים למעריך בכיר, אשר יפסוק אם החיבור ייחשב תגובה סבירה לגריין או לא. הערה: בחלק גדול מהמקרים שבהם חיבורים הוערכו כתגובה בלתי סבירה למטלה, הסיבה לכך הייתה שכלל לא נכתב חיבור. סביר שהדבר נובע מהעובדה שהבחינה לא הייתה בחינת אמת, ושהנבחנים שנמנעו מכתובה לא היו עושים זאת בבחינה תפעולית – ומכאן ששיעור החיבורים שיוערכו כתגובה בלתי סבירה למטלה צפוי להיות נמוך יותר, וכך גם שיעור ההסכמה בין שני המעריכים בנוגע למתן הציון 0.

פרסום המחוון (בחברת ההסבר ובאתר המרכז)

חברי צוות הפיתוח התלבטו אם כדאי לפרסם את המחוון. מחד גיסא, הפרסום עשוי להפחית את החרדה מפני מטלת הכתיבה, לאפשר לנבחנים הכנה ממוקדת וטובה יותר ולשדר שקיפות, הגינות וביטחון בכלי המדידה. מאידך גיסא, סביר שיהיו במחוון שינויים נוספים בעקבות תחילת השימוש התפעולי בו, ופרסומו בחוברת ההסבר עלול להגביל הכנסת שינויים כאלה. בסופו של דבר הוחלט לפרסם את המחוון בחוברת ההסבר ובאתר, בין היתר מתוך הנחה שתוכנו ידלוף לציבור ממילא, בגרסה מדויקת או מסולפת, מהמעריכים הרבים שייחשפו אליו. בחוברת ההסבר מצוין שהנוסח המדויק של המחוון המפורסם אינו מחייב, ובאתר תפורסם הגרסה העדכנית של המחוון לאחר כל שינוי בו.

עבודה ללא העברה ניסיונית

בדיקת המטלות בפרקי פיילוט תביא לחשיפתן, ולכן, שלא כפריטי הבררה במכפ"ל, המטלות יועברו בפרקים תפעוליים ללא בדיקה מקדימה. מצב זה בעייתי, שכן לעתים מתגלים פערים בין כוונות מנסחי המטלות ובין הבנת הנבחנים אותן, כפי שהתברר בשני הניסויים שנערכו במהלך פיתוח מטלת הכתיבה. כדי להתמודד עם בעיה זו נשקלת האפשרות להיעזר בכותבי חיבורים מכל רמות היכולת, שיכתבו חיבורים בתנאי אמת וישיבו על שאלוני משוב לכל מטלה שתפותח. קיים כמובן חשש מדליפת חומר, אולם ייתכן שיהיה אפשר למנוע זאת באמצעות הגבלת ההשתתפות בפיילוט לבעלי מחויבות למאל"ו בלבד (למשל עובדים בפרויקטים אחרים). ככל שיצטברו מטלות במאגר, כן תפחת סכנת החשיפה.

סיכום

תהליך פיתוח מטלת הכתיבה שתואר בדוח זה נמשך כשנתיים, מאוקטובר 2010 ועד אוקטובר 2012 – מועד ההעברה התפעולית הראשונה של המטלה. בתהליך זה נסקרו מבחני כתיבה מקבילים למכפ"ל שנמצאים בשימוש תפעולי, פותחו כמה חלופות לסוג המטלה ולמחווון שישמש להערכתה ותוכנן הליך מיון והכשרה למעריכי החיבורים. בתקופת הפיתוח נערכו שני ניסויים רחבי היקף, שסייעו בקבלת ההחלטות הנוגעות לסוג המטלה ומאפייניה, לאופי המחווון ושיטת הנקידה, ולהליך הכשרת המעריכים.

תוצאות הניסויים הביאו להתמקדות בכתיבת מטלות הדורשות כתיבה עיונית-טיעונית, ולפיתוח מחווון אנליטי דו-ממדי (תוכן ולשון). הוחלט שכל חיבור יוערך על ידי שני מעריכים בלתי תלויים, ושסולם ההערכה בכל אחד מהממדים ינוע בין 1 (חלש מאוד) ל-6 (מעולה). אחרי קבלת ההחלטות הנוגעות לאופי המטלות, המחווון ומודל ההערכה, נבנה בנק מטלות כתיבה תפעוליות.

הליך המיון וההכשרה של מעריכי החיבורים פותח אף הוא לפי אופי המטלות והמחווון. המיון מבוסס על עמידה בדרישות סף ראשוניות, ומבחן הכולל הערכה של כמה חיבורים. המעריכים המתלמדים עוברים סדנה לפני כל מועד הערכה, שכוללת הצגה של עקרונות ההערכה, אפיון המטלה הספציפית שתוערך והתנסות מעשית בהערכת חיבורים שנכתבו בתגובה למטלה זו. בנוסף פותחה על בסיס הסדנה הפרונטלית גם סדנה מתוקשבת, המיועדת למעריכים ותיקים.

ביבליוגרפיה

- גנאור-שטרן, ד', (1996), "סקירת ספרות בנושא מבחנים להערכת יכולת כתיבה". דוח מספר 214. מרכז ארצי לבחינות ולהערכה.
- גנאור-שטרן, ד', ריזל, ר' (1997), "פרוייקט מבחן כתיבה – סיכום ביניים". דוח פנימי. מרכז ארצי לבחינות ולהערכה.
- רוזנר, ר', (1995), "דגמים שונים לכתיבה עיונית". במכללה, 6, 17-27.
- Blake-Yancey, K. (1999). "Looking Back as We Look Forward: Historicizing Writing Assessment". *College Composition and Communication*; 50, 3, 483-503.
- Bloom, D. & Green J. (1984). "Directions in the Sociolinguistic Study of Reading". In P. D. Pearson (Ed), *Handbook of Reading Research*. New York & London: Longman; 395-421.
- Breland, H., Camp, R., Jones, R., Morris, M. & Rock, D. (1987). "Assessing Writing Skill". New York: College Entrance Examination Board.
- Breland, H., Bridgeman, B. & Fowles, M. (1999). "Writing Assessment in Admission to Higher Education: Review and Framework". RR No. 99-3. CollegeBoard.
- Crowhurst, M. & Piche, G. L. (1979). "Audience and Mode of Discourse Effects on Syntactic Complexity in Writing on Two Grade Levels". *Research in Teaching of English*, 13, 101-109.
- De Beaugrande, R. (1980). *Text, Discourse, and Process*. Norwood, New Jersey: Ablex.
- Flower, L. & Hayes, J. R. (1981). "A Cognitive Process Theory of Writing". *College Composition and Communication*, 32, 365-387.
- Freedman, S. W. (1981). "Influences on Evaluators of Expository Essays: Beyond the Text". *Research in the Teaching of English*, 15, 3, 245-55.
- Hale, G. (1992). "Effects of Amount of Time Allowed on the Test of Written English". *TOEFL Research Report*, 39. (ETS RR No. 92-27). Princeton, New Jersey: Educational Testing service.
- Heubert, J. P. & Hauser, R. M (Eds), (1999). *High stakes: Testing for Tracking, Promotion, and Graduation*. Washington, DC: National Academy Press.
- Kobrin, J., Patterson, B., Shaw, E., Mattern, K. & Barbuti, S. (2008). "Validity of the SAT for Predicting First-Year College Grade Point Average". RR. No. 2008-5. CollegeBoard.
- Langer, J. (1984). "The Effects of Available Information on Responses to School Writing Tasks". *Research in the Teaching of English*, 18, 1, 27-44.
- Mattern, K., Camara, W. & Kobrin, J. (2007). "SAT Writing: An Overview of Research and Psychometrics to Date". RR. No. 32. CollegeBoard.
- Messick, S. (1989). "Meaning and Values in Test Validation: The Science and Ethics of Assessment". *Educational Research*, 18, 2, 5-11.

Milewski, G., Johnsen, D., Glazer, N. & Kubota, M. (2005). "A Survey to Evaluate the Alignment of the New SAT Writing and Critical Reading Sections to Curricula and Instructional Practices". (ETS RR No. 05-07). Princeton, New Jersey: Educational Testing Service.

Murphy, S., Carroll, K., Kinzer, C. & Robyns, A. (1982). "A Study of the Construction of the Meaning of a Writing Prompt by its Authors, the Student Writers, and the Raters". In J. R. Gray & L. P. Ruth (Eds), *Properties of Writing Tests: A Study of Alternative Procedures for Holistic Writing Assessment* (386-468). Berkeley: University of California.

Murphy, S. (2003). "That was Then, This is Now: The Impact of Changing Assessment Policies on Teachers and the Teaching of Writing in California". *Journal of Writing Assessment*, 1, 1, 23-45.

Murphy, S. & Blake-Yancey, K. (2008). "Construct and Consequence: Validity in Writing Assessment". In C. Bazerman (Ed), *Handbook of Research on Writing: History, Society, School, Individual, Text* (365-385). London: Routledge.

Nelson, N. (2008). "The Reading-Writing Nexus in Discourse Research". In C. Bazerman (Ed), *Handbook of Research on Writing: History, Society, School, Individual, Text* (435-450). London: Routledge.

Oh, H. & Walker, M. (2006-7). "The Effects of Essay Placement and Prompt Type on Performance on the New SAT". RR No 2006-7. CollegeBoard.

Proctor, T. P. & Young Kounig, R. (2009). "Select Psychometric Properties and Predictive Validity of Scores on the SAT Writing Section". A Paper Presented at the Annual Meeting of the American Education Research Association San Diego.

Reid, J. (1990). "Responding to Different Topic Types: A Quantitative Analysis from a Contrastive Rhetoric Perspective". In B. Kroll (Ed), *Second Language Writing: Research Insights for the Classroom*. Cambridge: Cambridge University Press; 191-210

Sarig, G. & Folman, S. (1988). *Issues in Testing Academic Literacy: An Integrative Approach*. Tel Aviv: The Open University in Israel.

Shanahan, T. (1984). "Nature of the Reading-Writing Relation: An Exploratory Multivariate Analysis". *Journal of Educational Psychology*, 76, 466-477.

Sweedler-Brown, C. O. (1993). "ESL Essay Evaluation: The Influence of Sentence-Level and Rhetorical Features". *Journal of Second Language Writing*, 2, 1, 3-17.

Tardy, C. & Swales, J. (2008). "Form, Text, Organization, Genre, Coherence, and Cohesion". In C. Bazerman (Ed), *Handbook of Research on Writing: History, Society, School, Individual, Text* (565-581). London: Routledge.