
הוגנות מערכת המיון לאוניברסיטאות לפי שפת היבחנות במכפ"ל (ערבית או עברית) ומגדר

תמר קנת-כהן
אליוט טורוול
כרמל אורן

דצמבר 2011



דוח מרכז 376
ISBN:978-965-502-161-5

תוכן העניינים

2	תקציר
2	מבוא
3	שלוש חלופות להגדרת הטיה בבררה
5	ממצאי מחקרים אמפיריים – סקירה של הספרות
9	המחקר הנוכחי
10	שיטה
10	מדגם
11	משתני המחקר
11	עיבוד הנתונים
14	תוצאות
14	מחקר 1 – שפת היבחנות
18	מחקר 2 – מגדר
21	סיכום ודיון
26	מקורות
28	נספח 1 – חישוב המדד לגודל ההטיה על פי תנאי הגבול
30	נספח 2 – סטטיסטיים נצפים
32	נספח 3 – מתאמים בין ממדי ההטיה על פי שלושת המדדים
34	נספח 4 – בדיקת האפקט המשולב של שפת היבחנות ומגדר

תקציר

בעבודה זו נבדקה הוגנות מערכת המיון לאוניברסיטאות לפי שפת היבחנות במכפ"ל¹ (ערבית או עברית) ומגדר. במחקר 1 נבדקה ההוגנות בהתייחס לשפת היבחנות ובמחקר 2 נבדקה ההוגנות בהתייחס למגדר.

שאלת ההוגנות נבדקה משני היבטים: ההיבט של תוקף דיפרנציאלי, שבוחן קיומם של הבדלים בין קבוצות בתוקף הניבוי של החזאים, וההיבט של הטיה בבררה. הטיה בבררה נבדקה בשלוש גישות: מודל הרגרסיה של קלירי (Cleary, 1968) שבדק ניבוי דיפרנציאלי, תנאי גבול בבדיקת ניבוי דיפרנציאלי (Linn, 1984) והשוואת ממוצעים (Thorndike, 1971). במסגרת גישת תנאי הגבול, הוצע במחקר זה מדד לגודל ההטיה, שמשלים את המידע הנומינלי ("הטיה לרעה", "הטיה לטובה" או "לא נמצאה הטיה") שנאסף עד כה בעת יישום גישה זו.

המחקר בוצע על נתוני 57,717 סטודנטים מכל האוניברסיטאות בישראל, אשר למדו בשנה א' במחזורים 2004/5 עד 2007/8.

בבדיקת תוקף דיפרנציאלי במחקר 1 נמצא שתוקף הניבוי של כל החזאים גבוה יותר בקרב נבחני עברית. במחקר 2 נמצא יתרון קל, ולא עקבי מעבר לחזאים, לתקפיות בקרב נשים.

בבדיקת הטיה בבררה במחקר 1 נמצא שציון הסכם, החזאי המשמש למיון המועמדים, מוטה לטובת נבחני ערבית. הטיה זו לטובת נבחני ערבית היא תוצר של העובדה שהבגרות מוטה מאוד לטובתם ומכפ"ל אינו מוטה או מוטה פחות לטובתם (תלוי בהגדרה של הטיה בבררה). במחקר 2 נמצא שציון הסכם מוטה מעט לטובת נשים. הטיה קלה זו לטובת נשים היא תוצר של העובדה שהבגרות מוטה לטובתן ומכפ"ל מוטה לרעתן, אך מידת ההטיה שלו פחותה מזו של הבגרות.

במחקר נבדקה גם שאלת האינטראקציה בין שפת היבחנות לבין מגדר בממצאים של תוקף דיפרנציאלי והטיה בבררה. לא נמצאה אינטראקציה, במובן שכיוון הממצאים ביחס לכל אחד משני המשתנים – שפת היבחנות ומגדר – היה דומה ברמות השונות של המשתנה האחר. ממצאים מפורטים של הבדיקה מוצגים בנספח.

הדיון בממצאים כולל גם השוואה בין הגישות השונות לבדיקת הטיה בבררה, תוך ניסיון להסביר ולפקח חלקית על מקורות להבדלים ביניהן.

מבוא

בדיקות של הוגנות תהליך הבררה לאוניברסיטאות מתבצעות באופן שוטף במאלי"ו² ביחס לקבוצות אוכלוסייה שונות. בפרט, נעשו בעבר מחקרים אחדים לבדיקת הוגנות תהליך הבררה לפי שפת היבחנות (ערבית או עברית) במכפ"ל (ברון וגפני, 1989; ברונר, אללוף ואורן, 1996; ברונר, 2004; Turvall, Bronner, Kennet-Cohen, & Oren, 2008) ולפי מגדר (גפני ובלר, 1989; Azen, Bronner, & Gafni, 1999; Gafni, Beller, & Bronner, 2000).

¹ מבחן כניסה פסיכומטרי לאוניברסיטאות
² מרכז ארצי לבחינות ולהערכה

למחקר הנוכחי יש שתי מטרות: האחת, לספק נתונים מעודכנים בשאלת ההוגנות כלפי קבוצות שמוגדרות על פי משתני הרקע שפת היבחנות ומגדר, תוך התייחסות גם לקבוצות שמוגדרות על פי צירוף של שני המשתנים, דבר שנעשה בעבר באופן חלקי (Gafni et al., 2000), והשנייה, ליישם בו-זמנית ולהשוות בין שלוש גישות להגדרת הטיה בבררה: מודל הרגרסיה של קלירי (Cleary, 1968) שבדק ניבוי דיפרנציאלי, תנאי גבול בבדיקת ניבוי דיפרנציאלי (Linn, 1984) והשוואת ממוצעים ברוח מודל היחס הקבוע של תורנדייק (Thorndike, 1971). בהתייחס לגישה השנייה, תנאי גבול בבדיקת ניבוי דיפרנציאלי, מוצע במחקר הנוכחי מדד לגודל ההטיה, שמיועד להשלים את המידע הנומינלי ("הטיה לרעה", "הטיה לטובה" או "לא נמצאה הטיה") שנאסף עד כה בעת יישום גישה זו.

נפתח בתיאור של שלוש הגישות להגדרת הטיה בבררה. בהמשך ייסקרו ממצאי המחקרים שנעשו בעבר במאלי"ו בנושא של הוגנות כלפי נבחני ערבית ונשים, תוך התייחסות לגישות שנקטו לבדיקת הטיה בבררה. כמו כן יובא סיכום קצר של הגישות והממצאים בבדיקות מקבילות בארה"ב. פרק המבוא יחתום בתיאור קצר של המחקר המוצג.

שלוש חלופות להגדרת הטיה בבררה

בספרות הוצעו מודלים שונים, או הגדרות שונות, להטיה בבררה (לסקירה והשוואה בין המודלים ראו, למשל, Darlington, 1971; Hunter & Schmidt, 1976; Linn, 1973; Peterson & Novick, 1976). כפי שמציינים מרבית החוקרים שלעיל, השוני בין ההגדרות נובע משוני בשיפוטם ערכיים.

מודל הרגרסיה של קלירי (בדיקת ניבוי דיפרנציאלי)

ההגדרה המקובלת להטיה בבררה בספרות העולמית (ראו, למשל, Mattern, Patterson, Shaw, 2001; Kobrin, & Parbuti, 2008; Young, 2001) מבוססת על מודל הרגרסיה של קלירי (Cleary, 1968). מודל זה מגדיר הטיה בבררה כמצב של ניבוי דיפרנציאלי³ שבו שימוש בקו רגרסיה משותף לניבוי הקריטריון על ידי החזאי (כלי המיון) גורר ניבוי יתר או חסר⁴ של הקריטריון (ציוני שנה או באוניברסיטה, למשל). הטיה בבררה נבדקת על ידי השוואת קווי הרגרסיה של הקריטריון על החזאי בשתי קבוצות או על ידי חישוב הסטייה הממוצעת של ציון הקריטריון בפועל של חברי קבוצה מסוימת מציון הקריטריון המנובא מקו רגרסיה שמחושב מעבר לקבוצות (AERA, APA, & NCME, 2011, Standard 3.7).

³ כאן המקום להתייחס לטענתם של גמליאל וקאהן (2004) ש"הנחת השקילות בין הטיה בניבוי להטיה בבררה לא הוכחה אף פעם" (שם, עמ' 434, ההדגשה שלנו). על פי הבנתנו, עניין ההוכחה אינו רלוונטי כאן, מאחר שמדובר בבחירה ערכית להגדיר הטיה בבררה כניבוי דיפרנציאלי (הטיה בניבוי).

⁴ למען הדיוק, קלירי הדגישה בהגדרה של הטיה בבררה מצב של ניבוי חסר (Cleary, 1968, p. 115) של הקריטריון, ומאחר שסוגיית ההוגנות עסקה מן הסתם בקבוצות שהישגיהן בחזאי נמוכים, אפשר שנכון יותר להבין את קלירי כמציעה את מודל הרגרסיה להסקה על ניבוי חסר (ולא על ניבוי יתר) של הקריטריון בקבוצות אלה, מה שתואם את הרציונל של תנאי הגבול שהוצעו בהמשך. ניתן להבין את דרלינגטון (Darlington, 1971, p. 76) כתומך בפרשנות זו.

תנאי גבול (בבדיקת ניבוי דיפרנציאלי)

על פי גישת תנאי הגבול, הממצא הרווח של ניבוי יתר של הישגי קבוצות חלשות במחקרי הטיה המשתמשים במודל הרגרסיה של קלירי, לא מהווה עדות אמינה להטיה. כאמור, מודל הרגרסיה משווה בין קווי הרגרסיה לניבוי הקריטריון על ידי החזאי בשתי הקבוצות. לין (Linn, 1984) הראה, בהתבסס על ניתוח נתיבים שעשה בירנבאום (Birnbau, 1979, 1981), שכשמתייחסים לחזאי ולקריטריון כאל שני אינדיקטורים סימטריים של יכולת אמיתית חבויה (בלתי נצפית), אזי העובדה שאלה אינדיקטורים בני טעות (fallible) גורמת לכך שגם במצב "אמיתי" של חוסר הטיה (שמוגדר כמצב שבו לשויכות קבוצתית אין קשר ישיר עם החזאי או הקריטריון. כלומר, אם יש הבדלים בין הקבוצות בחזאי ובקריטריון, הם נובעים רק מהבדלים ביכולת האמיתית) יתקבל עם פי המודל של קלירי ממצא של ניבוי יתר של הקריטריון (הטיה לטובה) עבור הקבוצה החלשה כאשר מנבאים אותו על פי משוואת הרגרסיה של הקבוצה החזקה.

לפיכך, ממצא של ניבוי יתר של הקריטריון לקבוצה החלשה על סמך השוואת קווי הרגרסיה לניבוי הקריטריון על ידי החזאי אינו אמין. השוואה זו מאפשרת להסיק בביטחון רק על ניבוי חסר עבור קבוצה זו. כדי להסיק על ניבוי יתר יש להשוות בין קווי הרגרסיה ה"הפוכים" (לניבוי החזאי על ידי הקריטריון). כמו ברגרסיה ה"רגילה" (לניבוי הקריטריון על ידי החזאי), השוואה זו מאפשרת להסיק בביטחון רק על ניבוי חסר עבור הקבוצה החלשה, אך הפעם זה ניבוי חסר של החזאי, שפירושו ניבוי יתר של הקריטריון עבור קבוצה זו.

גישת תנאי הגבול הומלצה בעבר בספרות המקצועית (Cole & Moss, 1989), אך המחקר האמפירי בתחום המשיך להתנהל ברוח מודל הרגרסיה, והיא נזנחה בהמשך (Camili, 2006). אפשר שהעובדה שהממצאים שהניב מודל הרגרסיה היו נוחים לבוני מבחנים ולמשתמשים בהם, אחראית במידה מסוימת להתפתחות זו. בנוסף, התפיסה שהקריטריון והחזאי הם משתנים בעלי מעמד שווה בניתוח לא נראתה לגמרי מוצדקת בהקשר של בררה (Kane & Mroch, 2010). יש לציין חיסרון בולט של גישת תנאי הגבול, שהוא עוצמת המבחן הנמוכה שלה: רק אפקטים חזקים של הטיה מצליחים לחדור דרך המסננת המחמירה של גישה זו. שמרנות זו בהסקה על הטיה עשויה אף היא להיות נוחה, בדרכה, לממסד הבוחן או הממין, ובמובן זה, לעורר תהיות וביקורת.

השוואת ממוצעים (הפרש d-ים)

גישה זו נקשרת להמשגה האינטואיטיבית של מודלים שונים שמציעים בצורות שונות אילוצים על שיעור המתקבלים בכל קבוצה, ומגדירים באופן כללי הטיה בבררה כמצב שבו שיעור המתקבלים מקבוצה מסוימת הוא גבוה מדי או נמוך מדי בשל העובדה שהבררה נעשתה על בסיס החזאי ולא על בסיס הקריטריון. בפרט, גישה זו מתיישבת עם מודל היחס הקבוע של תורנדייק (Thorndike, 1971). תורנדייק ניסח את התנאי לחוסר הטיה בבררה כשוויון בין קבוצות במסך בין פרופורציית המועמדים מן הקבוצה שמתקבלים על סמך החזאי לבין הפרופורציה שהייתה מתקבלת על סמך הקריטריון. תנאי זה סובל מחוסר עקביות במצב של בררה לא הכרחית⁵ (Peterson & Novick,)

⁵ שבה מספר המועמדים אינו גדול ממספר המקומות, והבררה מונעת מהרצון לקבל רק מועמדים שהישיגיהם בקריטריון יהיו גבוהים מסף כלשהו

(1976), וכמענה לכך הציגו קאהן וגמליאל (Cahan & Gamliel, 2006) את מודל ההפרש הקבוע שעל פיו התנאי לחוסר הטיה בבררה הוא שוויון בין קבוצות *מפרש* בין פרופורציית המועמדים מן הקבוצה שמתקבלים על סמך החזאי לבין הפרופורציה שהייתה מתקבלת על סמך הקריטריון.

על פי שני המודלים, בררה הכרחית⁶ תהיה לא מוטה ביחס לקבוצה מסוימת אם הפרופורציה של המתקבלים מתוכה על סמך החזאי זהה לפרופורציה שהייתה מתקבלת על סמך הקריטריון. כאשר שיעור המתקבלים על סמך החזאי גבוה מן השיעור שהיה מתקבל על סמך הקריטריון נאמר שהחזאי מוטה לטובת חברי הקבוצה, ולהיפך, כאשר שיעור המתקבלים על סמך החזאי נמוך יותר, נאמר שהחזאי מוטה לרעת הקבוצה. במצב של שתי קבוצות משלימות, נגזר מהתנאי שלעיל, ובהנחות מסוימות לגבי החזאי והקריטריון (תוקף ניבוי חיובי מעבר לשתי הקבוצות, התפלגות דו-משתנית נורמלית לחזאי ולקריטריון בתוך כל קבוצה, עם שונות ותוקף ניבוי זהים בין הקבוצות), שהבררה אינה מוטה כאשר הפער בין שתי הקבוצות בחזאי זהה לפער ביניהן בקריטריון (גמליאל וקאהן, 2004; Darlington, 1971). ברוח זו תיבדק הטיה בבררה על ידי השוואת הפער (המתוקן) בין ממוצעי שתי הקבוצות בחזאי לפער ביניהן בקריטריון.

גישה זו זכתה לביקורת לא מבוטלת (נבון וכהן, 2009). הביקורת העיקרית כלפי גישה זו היא שהיא מתעלמת מטבעו הרגרסיבי של קו הניבוי של הקריטריון על ידי החזאי. על פי גישה זו, כל קבוצה שמזוהה עם הישגים נמוכים יחסית בחזאי (ובמקרה הקיצוני, מוגדרת על פי הישגים נמוכים בחזאי) תזוהה כסובלת מהטיה לרעה בתהליך הבררה. ניתן, כמובן, לטעון כנגד ביקורת זו, ששאלת ההוגנות צריכה להיבחן באופן בלתי תלוי במגבלת התוקף, ושאינן זה לגיטימי "להיעזר" בתוקף נמוך כדי להסיק על הטיה לטובת הקבוצה החלשה (במודל הרגרסיה) או כדי להימנע מהסקה (בגישת תנאי הגבול).

ממצאי מחקרים אמפיריים – סקירה של הספרות

מחקרי הוגנות בבררה שנעשו במאל"ו לפי שפת היבחנות (ערבית או עברית) במכפ"ל

המחקרים שנעשו במאל"ו לבדיקת הוגנות תהליך הבררה⁷ לפי שפת היבחנות (ערבית או עברית) במכפ"ל (ברון וגפני, 1989; ברונר ואחרים, 1996; ברונר, 2004; Turvall et al., 2008) משקפים תהליך התפתחותי בגישות ובשיטות שנקטו. סיכום של המאפיינים המתודולוגיים של המחקרים שנעשו מוצג להלן בלוח 1.

⁶ שבה הסיבה היחידה לדחיית מועמדים היא שמספרם גדול ממספר המקומות (המצב השכיח בהקשר של בררה להשכלה גבוהה)

⁷ אחדים מן המחקרים המדווחים כאן בדקו נושאים נוספים: ברון וגפני (1989) בדקו DIF, ברונר ואחרים (1996) בדקו הוגנות כפונקציה של שיטות שונות לכיול הציון המילולי בקרב נבחני ערבית, וברונר (2004) בדק את שאלת ההוגנות בהקשר של שימוש במצרף.

מחקרים לבדיקת הוגנות תהליך הבררה לפי שפת היבחנות (ערבית או עברית) במכפ"ל

מחקר	מחזורי המחקר ^a	מספר סטודי'		הטיה בבררה		תוקף דיפרנציאלי
		ערבית	ערבית	הגישה	תיקון ^b	
ברון וגפני, 1989	1983/84	2185	496	מודל הרגרסיה ^d	לא	לא ^c
ברונר ואחרים, 1996	1992/93 -1993/94	9999	1624	הגדרה I של דרלינגטון ^{e,d}	לא	כן ^f
				הגדרה II של דרלינגטון ^g	לא	
				הגדרה III של דרלינגטון ^h	לא	
ברונר, 2004	1992/93 -1997/98	29131	5373	ערכי d בחזאים ובקריטריון	כן	כן
Turvall et al., 2008	1995/96 -1998/99 2002/03	33784	7530	תנאי גבול ^d	כן	כן

^a תלמידי שנה א' בשנים המצוינות (להוציא את המחקר הראשון שבו השנים מתייחסות לשנות היבחנות במכפ"ל)
^b האם נעשה תיקון לקיצוץ תחום
^c האם נבדק תוקף דיפרנציאלי
^d יש התייחסות גם לערכי d בחזאים ובקריטריון : אצל ברון וגפני (1989) ואצל ברונר ואחרים (1996) ללא תיקון לקיצוץ תחום ; אצל טרוורל ואחרים (Turvall et al., 2008) עם תיקון.
^e (Darlington, 1971); יושם מודל הרגרסיה (Cleary, 1968) שמקביל להגדרה זו.
^f בהנחת בררה ישירה על ידי החזאי שאת התוקף שלו אומדים.
^g (Darlington, 1971); הגדרה זו משקפת את מודל היחס הקבוע של תורנדייק (Thorndike, 1971).
^h (Darlington, 1971); הגדרה זו היא מקרה פרטי של מודל ההסתברות המותנית של קול (Cole, 1973).

עיון בלוח 1 מלמד, בהתייחס לגישה בבדיקת הטיה בבררה, על מעבר בשלבים ממודל הרגרסיה של קלירי לתנאי הגבול לניבוי דיפרנציאלי: ההגדרות הראשונה והשלישית של דרלינגטון הן ניסוח חלופי (בהנחות מסוימות) של מודל הרגרסיה של קלירי ומודל ההסתברות המותנית של קול (Cole, 1973) בהתאמה, שמופיעות בהמשך כ"רגרסיה הרגילה" וה"רגרסיה ההפוכה" בהתאמה בתנאי הגבול. בדיקתן בנפרד (ברונר ואחרים, 1996) הובילה בהמשך לשילובן יחד בגישת תנאי הגבול (Turvall et al., 2008).⁸

השוואת ערכי d בחזאים ובקריטריון ליוותה את כל מחקרי ההוגנות, אם בהתייחסות מילולית בדיון בלבד (ברון וגפני, 1989) ואם כעיבוד סטטיסטי פורמלי (ברונר ואחרים, 1996; Turvall et al., 2008), אך רק במחקר אחד (ברונר, 2004) קיבלה השוואה זו מעמד של הגדרה להוגנות. השוואה זו מתיישבת עם מודל היחס הקבוע של תורנדייק (Thorndike, 1971).

עוד ניתן להיווכח שתיקון לקיצוץ תחום לא נכלל תמיד בבדיקת הטיה בבררה. הוא שולב לראשונה (בתחום של הוגנות כלפי נבחני ערבית) במחקר של ברונר (2004).

תוקף דיפרנציאלי נבדק לראשונה במחקר של ברונר ואחרים (1996), אם כי התיקון לקיצוץ תחום בעבודה זו אינו בר השוואה לתיקונים שנעשו בהמשך: הוא נעשה בהנחה (שאינה מתקיימת בפועל) שהבררה התבססה על החזאי שאת התוקף שלו מבקשים לאמוד.

הסיכום שלעיל מוביל למסקנה שקשה לעשות אינטגרציה של מחקרי העבר, שנבדלו זה מזה הן בגישות לבדיקת הוגנות והן בשיטות עיבוד הנתונים. עם זאת, ניתן לזהות עקביות בממצאים

⁸ גישת תנאי הגבול יושמה לראשונה במאליו במחקר על הוגנות כלפי נבחני מכפ"ל ברוסית (Gafni & Bronner, 1998).

שהתקבלו בעבודות שלעיל: בבדיקת הטיה בבררה מוצאים שציון הסכם מוטה לטובת נבחני ערבית. מידת ההטיה היא תלויה הגדרה, אך הכיוון נשמר בכולן. הבגרות מוטה מאוד לטובת נבחני ערבית, ובמכפ"ל הכיוון הוא תלוי הגדרה (אך ממדי ההטיה באופן כללי קטנים). בהתייחס למרכיבי מכפ"ל (התחום המילולי, התחום הכמותי ותחום האנגלית), כיוון ההטיה הוא תלוי הגדרה, אך מבין שלושתם התחום הכמותי מוטה יותר לטובת נבחני ערבית. לשם המחשה, ממצאי המחקר האחרון שלעיל (Turvall et al., 2008) הם כדלהלן: בבדיקת הטיה בבררה לפי תנאי הגבול נמצא שציון הסכם מוטה לטובת נבחני ערבית ב-21% מן החוגים⁹, הבגרות מוטה לטובתם ב-51% מן החוגים ומכפ"ל מוטה לרעתם ב-4% מן החוגים. בהתייחס למרכיבי מכפ"ל, התחום המילולי אינו מוטה, התחום הכמותי מוטה לטובת נבחני ערבית ב-7% מן החוגים, ואנגלית מוטה לרעתם ב-3% מן החוגים.

בבדיקת תוקף דיפרנציאלי מוצאים שתוקף כל החזאים גבוה יותר בקרב נבחני עברית (גם כאשר המתאמים מתוקנים לקיצוץ תחום באופן שמניח סטיית תקן זהה לחזאי הממייין בשתי הקבוצות).

מחקרי הוגנות בבררה שנעשו במאל"ו לפי מגדר

המחקרים שנעשו במאל"ו לבדיקת הוגנות תהליך הבררה לפי מגדר (גפני ובלר, 1989; Azen et al., 2000; Gafni et al., 2000¹⁰) משקפים תהליך התפתחותי מקביל למה שתואר לעיל ביחס לשפת היבחנות במכפ"ל. סיכום של המאפיינים המתודולוגיים של המחקרים שנעשו מוצג להלן בלוח 2.

לוח 2

מחקרים לבדיקת הוגנות תהליך הבררה לפי מגדר

מחקר	מחזורי המחקר ^a	מספר סטוד'ים		הטיה בבררה		תוקף דיפרנציאלי	
		גברים	נשים	הגישה	תיקון ^b	בדיקה ^c	תיקון ^b
גפני ובלר, 1989	1984/85	3069	3134	מודל הרגרסיה ^d	לא	כן	לא
Azen et al., 1999	1991/92 -1995/96	26278	35607	תנאי גבול ^d	לא	כן	לא
Gafni et al., 2000	1985/86 -1996/97	59580	75533	תנאי גבול ^d	לא	כן	כן

^a תלמידי שנה א' בשנים המצוינות

^b האם נעשה תיקון לקיצוץ תחום

^c האם נבדק תוקף דיפרנציאלי

^d יש התייחסות גם לערכי d בחזאים ובקריטריון, ללא תיקון לקיצוץ תחום.

עיון בלוח 2 מראה, גם במחקר על הוגנות לפי מגדר, על מעבר ממודל הרגרסיה של קלירי לתנאי הגבול בבדיקת הטיה בבררה. השוואת ערכי d בחזאים ובקריטריון ליוותה את כל המחקרים, אך לא במעמד של הגדרה להוגנות. לא נעשה תיקון לקיצוץ תחום, לא ביחס לסטטיסטיים של הטיה בבררה ולא ביחס לערכי ה-d.

תוקף דיפרנציאלי נבדק בכל המחקרים, אך רק במחקר האחרון תוקנו התקפוויות לקיצוץ תחום.

⁹ במקרים שבהם המגמה לא הייתה עקבית בין חוגים, חושב ההפרש בין אחוז החוגים עם הטיה בכיוון אחד לבין אחוז החוגים עם הטיה בכיוון השני (המספר הגדול פחות המספר הקטן). לדוגמה, בהתייחס לציון הסכם, נמצאה הטיה לטובת נבחני ערבית ב-22% מן החוגים והטיה לרעתם ב-1% מן החוגים. לכן מדווח כאן שיש הטיה לטובה ב-21% מן החוגים.

¹⁰ פורסם גם בכתב העת Applied Measurement in Education (Azen, Bronner, & Gafni, 2002).

המחקרים הראשון והשני בוצעו בקרב נבחני עברית במכפ"ל. המחקר השלישי בוצע בקרב כל אחת משלוש קבוצות של נבחני מכפ"ל: נבחני עברית, נבחני ערבית ונבחני רוסית. את רוח הממצאים שהתקבלו בעבודות שלעיל ניתן לסכם כך: בבדיקת הטיה בבררה מוצאים שציון הסכם אינו מוטה. בגרות ומכפ"ל פועלים בכיוונים מנוגדים: הבגרות מוטה לטובת נשים ומכפ"ל מוטה לרעתן. מבין שלושת מרכיבי מכפ"ל התחום הכמותי הוא המוטה ביותר לרעת נשים. כך למשל, ממצאי המחקר האחרון שלעיל (Gafni et al., 2000) הם כדלהלן: בבדיקת הטיה בבררה לפי תנאי הגבול בקרב נבחני עברית נמצא שציון הסכם מוטה לטובת נשים ב-2% מן החוגים¹¹. בקרב נבחני ערבית, הוא מוטה לטובת נשים ב-10% מן החוגים. הבגרות מוטה לטובתן ב-15% וב-13% מן החוגים בקרב נבחני עברית ונבחני ערבית בהתאמה. מכפ"ל מוטה לרעת נשים בקרב נבחני עברית ב-11% מן החוגים ולטובת נשים בקרב נבחני ערבית ב-2% מן החוגים. בהתייחס למרכיבי מכפ"ל, בשתי השפות התחום הכמותי הוא המוטה ביותר לרעת נשים – במידה רבה יותר בקרב נבחני עברית (14% מן החוגים) ובמידה פחותה (4% מן החוגים) בקרב נבחני ערבית. בבדיקת תוקף דיפרנציאלי מוצאים הבדלים קטנים בתקפויות בין גברים לנשים. במחקר האחרון נמצא, בשתי השפות, תוקף גבוה יותר בקרב נשים.

מחקרי הוגנות בבררה שנעשו ביחס ל-SAT¹²

כאמור, מחקרי ההוגנות של ה-SAT (סקירה נרחבת מופיעה אצל Young, 2001; מחקר עדכני, הבודק את הוגנות ה-SAT אחרי השינויים שהוכנסו בו במארכס 2005, מדווח אצל Mattern et al., 2008) בודקים הטיה בבררה על פי מודל הרגרסיה של קלירי, על ידי חישוב הסטייה הממוצעת של ציון הקריטריון בפועל של חברי קבוצה מסוימת מציון הקריטריון המנובא מקו רגרסיה שמחושב מעבר לקבוצות. לא מתבצע תיקון לקיצוץ תחום, למרות שהספרות (Linn, 1983) מציינת שאומדי הפרמטרים של משוואות ניבוי על בסיס חזאי שעבר בררה עקיפה עלולים להיות מוטים¹³. הממצאים מעידים על הטיה לטובת קבוצות מיעוט אתניות (סטודנטים היספנים ושחורים) על ידי כל החזאים (SAT ומרכיביו וממוצע ציוני בית ספר תיכון) וצירופיהם. תמונה הפוכה (הטיה לרעה) מתקבלת ביחס לנשים.

בבדיקת תוקף דיפרנציאלי מתבצע תיקון לקיצוץ תחום¹⁴. הממצאים מעידים על תוקף ניבוי נמוך יותר בקרב קבוצות מיעוט אתניות (בהשוואה לתוקף בקרב סטודנטים לבנים) ותוקף גבוה יותר בקרב נשים (בהשוואה לתוקף בקרב גברים).

¹¹ במקרים שבהם המגמה לא הייתה עקבית בין חוגים, מדווח ההפרש בין אחוז החוגים עם הטיה בכיוון אחד לבין אחוז החוגים עם הטיה בכיוון השני (המספר הגדול פחות המספר הקטן).

¹² מבחן ההערכה האקדמית (Scholastic Assessment Test) המשמש לבררת מועמדים ללימודים בקולגים בארה"ב
¹³ הכיוון הטיפוסי הוא השטחת קווי הרגרסיה, מה שמעלה את הנטייה לקבל ממצא של הטיה לטובת הקבוצה הנמוכה בחזאי.

¹⁴ התיקון הרב-משתני של Pearson-Lawley, כאשר האוכלוסייה היא המחזור הרלוונטי של תלמידי תיכון שמיועדים ללמוד בקולג' (Mattern et al., 2008).

המחקר הנוכחי

המחקר הנוכחי בדק את הוגנות מערכת המיון לאוניברסיטאות כלפי נבחני מכפ"ל בערבית וכלפי נשים. באופן מפורט, הוא כלל שני מחקרים:

מחקר 1 – הוגנות לפי שפת היבחנות – שכולל שתי קבוצות: נבחני עברית ונבחני ערבית (מעבר למגדר);

מחקר 2 – הוגנות לפי מגדר – שכולל שתי קבוצות: גברים ונשים (מעבר לשפת היבחנות); שאלת ההוגנות נבדקה משני היבטים: ההיבט של תוקף דיפרנציאלי, שבוחן קיומם של הבדלים בין קבוצות בתוקף הניבוי של החזאים, וההיבט של הטיה בבררה. שלוש שיטות לבדיקת הטיה בבררה יושמו במחקר הנוכחי: מודל הרגרסיה של קלירי (Cleary, 1968) שבודק ניבוי דיפרנציאלי, תנאי גבול בבדיקת ניבוי דיפרנציאלי (Linn, 1984) והשוואת ממוצעים (Thorndike, 1971). שילובן של שלוש השיטות במחקר אחד מהווה נדבך נוסף בהתפתחות החשיבה בשאלת ההמשגה של הטיה בבררה. כפי שתואר בלוחות 1 ו-2 שלעיל, מודל הרגרסיה של קלירי, שהופיע במחקרי ההוגנות המוקדמים, הוחלף בהדרגה בגישת תנאי הגבול. מסיבות שאליהן התייחסו בחלקים מוקדמים יותר של המבוא, מצאנו לנכון לחזור ולשלב מודל זה בבדיקת הטיה בבררה. במקביל, נכללה גם השוואת ממוצעים (הפרש d-ים), שכפי שצוין בהתייחס ללוחות 1 ו-2 שלעיל, ליוותה את כל מחקרי ההוגנות שבוצעו עד כה, לעיתים קרובות כסרח עודף שמעמדו יחסית להגדרה ששימשה במחקר באופן רשמי, אינו ברור. גישת תנאי הגבול מאזנת במידה מסוימת את גישת השוואת הממוצעים, בכך שהיא מביאה בחשבון את תוקף הניבוי של כלי המיון כאשר בודקים את הוגנותו (אך, להבדיל ממודל הרגרסיה, היא עושה זאת באופן סימטרי ביחס לשני כיווני הטיה).

עבור בדיקת הטיה בבררה על פי תנאי גבול לניבוי דיפרנציאלי הוספנו במחקר הנוכחי מדד חדש לגודל ההטיה. מדד זה משלים את המידע הנומינלי ("הטיה לרעה", "הטיה לטובה" או "לא נמצאה הטיה") שנגזר עד כה בעת יישום גישה זו. מדד זה מחושב ברוח תנאי הגבול, כלומר: על פי משוואת הרגרסיה הרלוונטית להכרעה על קיומה של הטיה. מדד זה מחושב תמיד (ולא משנה מהי משוואת הרגרסיה "קובעת": "ה"רגילה" או "ה"פוכה") כמרחק האנכי (כלומר, המרחק במונחי הקריטריון) בין קו הרגרסיה של קבוצת המוקד לקו הרגרסיה של קבוצת ההתייחסות (כאשר הציון בחזאי שווה לציון הממוצע בחוג).

מן הראוי להפנות את תשומת הלב להחלטה שהתקבלה במחקר הנוכחי, ליישם את הגישה הראשונה להגדרת הטיה בבררה – מודל הרגרסיה של קלירי – ברוח ההמלצות המקצועיות (AERA et al., 2011)¹⁵ והמחקר האמפירי בתחום (Mattern et al., 2008), כלומר: מבלי לתקן לקיצוץ תחום. החלטה זו מתיישבת עם הרצון להציג את הממצאים כפי שהיו מתקבלים לו יושם המודל המומלץ והרווח להגדרת הטיה בבררה. כמובן שבהשוואה בין הממצאים שהתקבלו בגישה זו לבין אלו שהתקבלו בשתי הגישות האחרונות (שבהן נעשה תיקון לקיצוץ תחום) הובא בחשבון הבדל זה ביישום, כפי שיוצג בהמשך.

¹⁵ נשלחה הערה על טיוטת התקנים למבחנים חינוכיים ופסיכולוגיים בעניין זה.

בנספח למחקר מוצגים ממצאים מבדיקה של האפקט המשולב של שני משתני הרקע¹⁶ – שפת היבחנות ומגדר. לצורך זה הוגדרו ארבע קבוצות: גברים נבחני עברית, גברים נבחני ערבית, נשים נבחנות עברית ונשים נבחנות ערבית. כמו במחקרים 1 ו-2 נבדק גם כאן ההיבט של תוקף דיפרנציאלי וההיבט של הטיה בבררה. ההיבט של הטיה בבררה נבדק רק באמצעות מודל הרגרסיה של קלירי, שהוא מתאים במיוחד לטיפול במצב של יותר משתי קבוצות (כאשר הוא נבדק על ידי חישוב הסטייה הממוצעת של ציון הקריטריון בפועל של חברי קבוצה מסוימת מציון הקריטריון המנובא מקו רגרסיה שמחושב מעבר לכל קבוצות), ומציב אילוצים מינימליים על מספר התצפיות מכל קבוצה.

ההוגנות נבדקה ביחס לכל אחד מכלי המיון המהווים חלק מתהליך הקבלה ללימודים לתואר ראשון באוניברסיטאות בישראל: בגרות, מכפ"ל ומרכיביו, וציון הסכם המבוסס על בגרות ומכפ"ל במשקלות שווים.

המחקר בוצע על נתוני תלמידי שנה א' בארבעה מחוזים בכל האוניברסיטאות בישראל.

שיטה

מדגם

המדגם הכללי שממנו נגזרו מדגמי שני המחקרים כלל 75,064 רשומות של סטודנטים מכל האוניברסיטאות בישראל, אשר למדו בשנה א' במחזורים תשס"ה (2004/5) עד תשס"ח (2007/8) ב-1,142 יחידות עיבוד, כאשר יחידת עיבוד מוגדרת כחוג לימוד במוסד ובמחזור (בהמשך נשתמש, לשם פשטות, במונח "סטודנט" לתיאור "רשומה של סטודנט" ובמונח "חוג" לתיאור "יחידת עיבוד").

סטודנט נכלל במדגם הראשוני אם:

- נבחן במכפ"ל בעברית או בערבית במבחן הקובע¹⁷;
- דיווח על המגדר שלו בהיבחנות אחת לפחות¹⁸;
- נתונים ציוניו בקריטריון ובכל החזאים (יפורטו בהמשך);

מדגמי שני המחקרים התבססו, בנוסף לתנאים שלעיל, גם על תנאי להיכללות חוג:

במחקר 1 (שפת היבחנות): לפחות 5 סטודנטים מכל שפת היבחנות

במחקר 2 (מגדר): לפחות 5 סטודנטים מכל מגדר

רק חוגים שניתן לסווגם כמילוליים או כמותיים¹⁹ נכללו במחקר. לא נכללו מסלולים למצטיינים או לימודים כלליים.

¹⁶ אפקט שזכה להתייחסות חלקית בעבר במחקר שבדק הוגנות כלפי קבוצות שמוגדרות לפי מגדר בתוך שפת היבחנות (Gafni et al., 2000)

¹⁷ המבחן שבו קיבל הסטודנט את הציון הכללי הגבוה ביותר לפני תחילת הלימודים

¹⁸ עבור סטודנטים שנבחנו יותר מפעם אחת דרשנו שלא תהיה סתירה בדיווח בין היבחנויות. כל ההיבחנויות של הסטודנט עד דצמבר 2010 הובאו בחשבון בעניין זה.

¹⁹ תחומי לימוד מילוליים כוללים את הפקולטות רוח ואמנויות, חברה - חוגים מילוליים (למשל, סוציולוגיה, פסיכולוגיה) ומשפטים; תחומי לימוד כמותיים כוללים את הפקולטות טבע, הנדסה וארכיטקטורה, חברה - חוגים כמותיים (למשל, כלכלה, סטטיסטיקה), רפואה ועזר-רפואה.

משתני המחקר

קריטריון

ציון שנה א' (צש"א), שמוגדר בטווח 10-100. ציון זה הוגדר בחוג כממוצע משוקלל בנקודות זכות (4 לפחות) של ציוני הקורסים אותם למד הסטודנט בשנה א' לתואר בוגר.

חזאים

1. ציון סכס, אשר חושב כצירוף במשקלות שווים של מכפ"ל ובגרות בקרב המועמדים לכל מוסד, עם ממוצע 50 וסטיית תקן 10 (בקרב המועמדים הנ"ל)²⁰.
2. ממוצע ציוני תעודת הבגרות (בגרות), כפי שמחושב ע"י האוניברסיטאות (כולל בונוסים).
3. הציון הכללי במכפ"ל במבחן הקובע. על פי מבחן זה נקבעו גם הציונים בתחומי מכפ"ל (חזאים 4 עד 6 להלן).
4. הציון בתחום חשיבה מילולית במכפ"ל (מילולי).
5. הציון בתחום חשיבה כמותית במכפ"ל (כמותי).
6. הציון בתחום אנגלית במכפ"ל (אנגלית).

עיבוד הנתונים

בכל חוג נעשו העיבודים הבאים:

סטטיסטיים תיאוריים

חושבו ממוצע וסטיית תקן של משתני המחקר בכל קבוצה.

תוקף דיפרנציאלי

תוקף הניבוי של כל אחד מן החזאים בכל קבוצה חושב על ידי מקדם מתאם פירסון בין החזאי לבין הקריטריון. התיקון לקיצוץ תחום נעשה עבור בררה חד-משתנית (על פי ציון קבלה), במקרה התלת-משתני (Gulliksen, 1950). תיקון זה מניח שידועה סטיית התקן של ציון הקבלה (ציון הסכס) ברמת המועמדים. כאומד לסטיית התקן של ציון הסכס בקרב המועמדים חושב שורש של ממוצע משוקלל של שונות ציון הסכס בקרב המועמדים לחוגי הלימוד במחזורים תשנ"ב ותשנ"ג. נכללו בחישוב רק מועמדים שנבחנו במכפ"ל בעברית או בערבית, שיש דיווח על המגדר שלהם ושיש להם נתונים בכל החזאים, ובתנאי שבחוג יש לפחות 20 מועמדים שעומדים בדרישות הנ"ל²¹. חשוב להדגיש שתקפויות הניבוי בשתי הקבוצות תוקנו לערך זהה של סטיית התקן של ציון הסכס, וזאת במטרה לנטרל הבדלי תוקף שמקורם בהבדלים בין הקבוצות בשונות המשתנים.

²⁰ באופן מפורט, ציון הסכס חושב באופן הבא: 1. חושבו ממוצע וסטיית תקן של מכפ"ל ובגרות בקרב המועמדים למחזור תשנ"ב ולמחזור תשנ"ג בכל מוסד בנפרד. 2. חושבו ממוצעים, מעבר לשני המחזורים (משוקללים במספר המועמדים לכל מחזור) של הסטטיסטיים הנ"ל בכל מוסד. 3. ציוני מכפ"ל ובגרות (של הלומדים ושל המועמדים) תוקנו לפי הממוצעים שחושבו בסעיף 2. 4. ציון סכס גולמי חושב בקרב הלומדים ובקרב המועמדים כסכום הציונים המתוקנים של מכפ"ל ובגרות. 5. ציון סכס מתוקן חושב כמו שמתואר בסעיפים 1 עד 3 ביחס למכפ"ל ולבגרות. 6. ציון הסכס המתוקן (של המועמדים ושל הלומדים) הוכפל ב-10 והוסף ל-50.
²¹ 456 חוגים נכללו בחישוב (לעומת 461 שהיו נכללים ללא הדרשה על שפת היבחנות של עברית או ערבית ודיווח על המגדר).

הטיה בבררה

על פי מודל הרגרסיה של קלירי ("קלירי")

ציון הקריטריון תוקנן על פי הממוצע וסטיית התקן מעבר לכל הסטודנטים בחוג. התקנון בוצע כדי להביא את ציוני הקריטריון לסולם בר השוואה בין חוגים (כיוון שהתוצאות המוצגות הן ממוצעים מעבר לחוגים).

עבור כל אחד מן החזאים חושבה משוואת רגרסיה לניבוי הקריטריון מעבר לכל הלומדים בחוג. עבור כל קבוצה חושבה הסטייה הממוצעת בין ציון הקריטריון המנובא על פי משוואת הרגרסיה המשותפת לבין הציון בפועל (ציון מנובא פחות ציון בפועל).

ערך חיובי מבטא הטיה לטובת הקבוצה וערך שלילי הטיה לרעה.

על פי תנאי גבול

בכל מחקר הוגדרה קבוצת מוקד (במחקר 1: נבחני ערבית; במחקר 2: נשים) וקבוצת התייחסות (במחקר 1: נבחני עברית; במחקר 2: גברים). בהתאם, הוגדר משתנה של שיוך קבוצתי שערכיו הם: 0 = קבוצת המוקד, 1 = קבוצת התייחסות. קבוצת המוקד תכונה להלן "חלשה יותר" אם ממוצע ציוניה בחזאי נמוך מממוצע ציוני קבוצת התייחסות, ו"חזקה יותר" אם ממוצע ציוניה בחזאי גבוה מממוצע ציוני קבוצת התייחסות.

בשלב הבא חושבו שתי משוואות רגרסיה: הראשונה, לניבוי הקריטריון על ידי החזאי, המשתנה של שיוך קבוצתי, ואינטראקציה בין החזאי לבין שיוך קבוצתי; והשנייה, לניבוי החזאי על ידי הקריטריון, המשתנה של שיוך קבוצתי ואינטראקציה בין הקריטריון לבין שיוך קבוצתי.

הסקה על ניבוי יתר (הטיה לטובה) או ניבוי חסר (הטיה לרעה) של ציוני הקריטריון של קבוצת המוקד התקבלה, באופן עקרוני, על פי כיוונו של המקדם של משתנה השיוך הקבוצתי²². השאלה איזה משוואת רגרסיה (ראשונה או שנייה) ואיזה כיוון של המקדם (חיובי או שלילי) משמשים בהסקה זו תלויה בשאלה אם קבוצת המוקד חלשה יותר (המצב הטיפוסי) או חזקה יותר, כפי שיפורט להלן.

הסקה על הטיה לרעת קבוצת המוקד התקבלה אם קבוצת המוקד חלשה יותר והמקדם של משתנה השיוך הקבוצתי במשוואת הרגרסיה הראשונה הוא שלילי, או אם קבוצת המוקד חזקה יותר והמקדם של משתנה השיוך הקבוצתי במשוואת הרגרסיה השנייה הוא חיובי. הסקה על הטיה לטובת קבוצת המוקד התקבלה אם קבוצת המוקד חזקה יותר והמקדם של משתנה השיוך הקבוצתי במשוואת הרגרסיה הראשונה הוא חיובי, או אם קבוצת המוקד חלשה יותר והמקדם של משתנה השיוך הקבוצתי במשוואת הרגרסיה השנייה הוא שלילי (פירוט נוסף של הרציונל של תנאי הגבול ושל אופן יישומם מצוי אצל קנת-כהן, 2001).

בנוסף למסקנה הנומינלית ("הטיה לרעה", "הטיה לטובה" או "לא נמצאה הטיה") חושב מדד לגודל ההטיה. מדד זה חושב, על פי משוואת הרגרסיה שממנה התקבלה המסקנה על הטיה, כהפרש בין ציון הקריטריון (כאשר הציון בחזאי שווה לציון הממוצע בחוג) על פי קו הרגרסיה של

²² ובתנאי שהמקדם של משתנה האינטראקציה קטן בערכו המוחלט מהמקדם של משתנה השיוך הקבוצתי (מה שמבטיח שבטווח של \pm סטיית תקן מן הממוצע הכללי בחזאי אין היפוך בסדר קווי הרגרסיה של שתי הקבוצות)

קבוצת ההתייחסות לבין ציון הקריטריון (במצב הנ"ל) על פי קו הרגרסיה של קבוצת המוקד. ערך חיובי של הפרש זה מבטא הטיה לטובת קבוצת המוקד וערך שלילי הטיה לרעתה. חוגים שבהם לא ניתן היה לזהות הטיה קיבלו את הערך 0 בממד זה. זהו, כאמור, מדד חדש. הרציונל שלו ואופן חישובו מוצגים בנספח 1.

שתי משוואות הרגרסיה חושבו על בסיס מטריצה של מתאמי פירסון (בין ארבעת המשתנים הרלוונטיים: קריטריון, חזאי, שיוך קבוצתי ואינטראקציה בין הקריטריון או בין החזאי לבין שיוך קבוצתי) שתוקנו לקיצוץ תחום.

מאחר שמשוואות הרגרסיה חושבו על בסיס מטריצה של מתאמים ולא ישירות מתוך הערכים במשתנים, המדד לגודל ההטיה (ההפרש בין ציוני הקריטריון בשתי המשוואות) הוא במונחי סטיות תקן, כמו שני המדדים האחרים לגודל ההטיה.

על פי פער מתוקן בין ממוצעי הקבוצות בקריטריון פחות הפער בחזאי ("הפרש d-ים")

המדד Cohen's d לגודל אפקט הוא הפער המתוקן בין ממוצעי שתי קבוצות במשתנה רציף. d הוגדר כממוצע קבוצת ההתייחסות פחות ממוצע קבוצת המוקד (חלקי ממוצע סטיות התקן) במשתנה הנדון. המדד להטיה בבררה הוא ה-d בקריטריון פחות ה-d בחזאי. ערך חיובי של הפרש זה מבטא הטיה לטובת קבוצת המוקד וערך שלילי הטיה לרעתה. במחקר הנוכחי חושב d כפונקציה של מקדם המתאם של פירסון בין משתנה השיוך הקבוצתי לבין המשתנה הנדון (הקריטריון או אחד מן החזאים) ושל הפרופורציות של כל אחת משתי הקבוצות בצירוף של שתיהן על פי הנוסחה²³: $d = r / \sqrt{pq(1 - r^2)}$, כאשר r הוא מקדם המתאם פירסון בין משתנה השיוך הקבוצתי (0 = קבוצת המוקד, 1 = קבוצת ההתייחסות) לבין המשתנה הנדון ו-p ו-q הם הפרופורציות של כל אחת משתי הקבוצות בצירוף של שתיהן. לצורך חישוב הערך המתוקן לקיצוץ תחום של d השתמשנו ב-r המתוקן לקיצוץ תחום וב-p ו-q בקרב הלומדים בחוג.

בפרק התוצאות יוצגו ממוצעים משוקללים (במספר הכולל של סטודנטים בחוג) של הסטטיסטיים השונים לפי תחומי לימוד (מילוליים או כמותיים) ומעבר לכולם. מספר הסטודנטים לפי תחומי לימוד וקבוצה ומספר החוגים לפי תחומי לימוד עבור כל אחד משני המחקרים מוצגים בלוח 3.

²³ ראו פיתוח של הנוסחה, בהתבסס על כהן (Cohen, 1988) אצל קנת-כהן (2001).

לוח 3

מספר הסטודנטים ומספר החוגים בשני המחקרים

סה"כ	תחומי לימוד		
	מילוליים	כמותיים	
מחקר 1 (שפת היבחנות)			
30561	15382	15179	סטודנטים עברית
8009	3709	4300	סטודנטים ערבית
376	206	170	חוגים
מחקר 2 (מגדר)			
31428	19355	12073	סטודנטים גברים
38853	16321	22532	סטודנטים נשים
833	441	392	חוגים

תוצאות

מחקר 1 – שפת היבחנות

בלוח 4 מוצגים סטטיסטיים תיאוריים של משתני המחקר.

לוח 4

ממוצעים וסטייות תקן של החזאים והקריטריון לפי שפת היבחנות ותחומי לימוד

תחומי לימוד	משתנה	עברית		ערבית	
		ממוצע	סטיית תקן	ממוצע	סטיית תקן
מילוליים	צש"א	84.5	7.2	73.7	9.8
	סכס	56.0	7.0	55.3	6.2
	בגרות	96.0	6.9	102.1	6.7
	מכפ"ל	599.4	63.4	532.8	50.9
	מילולי	118.3	13.4	106.7	12.5
	כמותי	115.7	13.2	108.9	11.9
	אנגלית	119.2	16.1	97.0	13.3
כמותיים	צש"א	80.5	9.7	70.7	11.4
	סכס	62.3	5.0	62.9	3.9
	בגרות	103.0	5.8	109.9	4.0
	מכפ"ל	668.3	42.7	617.4	39.4
	מילולי	127.7	10.5	119.9	11.3
	כמותי	131.0	9.1	127.4	9.6
	אנגלית	130.5	12.4	108.4	14.0
כולם	צש"א	82.5	8.5	72.2	10.6
	סכס	59.1	6.1	59.1	5.1
	בגרות	99.5	6.4	106.0	5.6
	מכפ"ל	633.5	54.2	574.6	45.6
	מילולי	122.9	12.1	113.2	11.9
	כמותי	123.2	11.4	118.1	10.8
	אנגלית	124.8	14.4	102.6	13.6

הממצאים המוצגים בלוח 4 מעידים שנבחני עברית גבוהים מנבחני ערבית במכפ"ל בכסטיית תקן אחת. פער דומה בגודלו אך הפוך בכיוונו מתקיים בבגרות. כתוצאה מכך דומים ממוצעי ציון הסכס בשתי הקבוצות, וזאת לעומת ציון שנה א' שבו יש פער של כסטיית תקן אחת (זהה לפער

במכפ"ל) לטובת נבחני עברית. מבין תחומי מכפ"ל הפער בין הקבוצות (לטובת נבחני עברית) הוא הגדול ביותר באנגלית והקטן ביותר בתחום הכמותי.

אין הבדלים משמעותיים בין תחומי הלימוד בעניין זה.

בלוח 5 מוצגים מקדמי המתאם בין החזאים לקריטריון.

לוח 5

מקדמי התוקף של החזאים לפי שפת היבחנות ותחומי לימוד²⁴

ערבית	עברית	חזאי	תחומי לימוד
0.37	0.51	סכם	מילוליים
0.31	0.46	בגרות	
0.33	0.42	מכפ"ל	
0.25	0.38	מילולי	
0.26	0.35	כמותי	
0.24	0.28	אנגלית	
0.39	0.53	סכם	כמותיים
0.36	0.50	בגרות	
0.32	0.41	מכפ"ל	
0.19	0.31	מילולי	
0.28	0.39	כמותי	
0.27	0.26	אנגלית	
0.38	0.52	סכם	כולם
0.33	0.48	בגרות	
0.32	0.41	מכפ"ל	
0.22	0.35	מילולי	
0.27	0.37	כמותי	
0.26	0.27	אנגלית	

הממצאים המוצגים בלוח 5 מעידים שתוקף הניבוי של ציון הסכם ושני מרכיביו העיקריים גבוה יותר בקרב נבחני עברית. מבין תחומי מכפ"ל הפער לטובת נבחני עברית מתקיים בתחום המילולי ובתחום הכמותי. תוקף הניבוי של אנגלית דומה בשתי הקבוצות.

אין הבדלים משמעותיים בין תחומי הלימוד בעניין זה.

בלוח 6 מוצגים שלושת המדדים להטיה בבררה כלפי נבחני עברית.

²⁴ מתוקנים לקיצוץ תחום. מקדמי תוקף נציפים מוצגים בלוח 12 בנספח 2.

הפרש d-ים			תנאי גבול				קלירי ^a	חזאי	תחומי לימוד
d קריטי פחות חזאי ^a	d חזאי ^b	d קריטי ^b	גודל ההטיה ^a	% חוגים שבהם					
				הטיה לרעה	אין הטיה	הטיה לטובה			
1.44	-0.09	1.35	0.18	0	65	35	1.00	סכס	מילוליים
2.33	-0.98	1.35	0.47	0	35	65	1.08	בגרות	
0.42	0.93	1.35	0.00	1	98	1	0.71	מכפ"ל	
0.57	0.78	1.35	0.01	1	95	4	0.80	מילולי	
0.96	0.39	1.35	0.04	1	88	11	0.92	כמותי	
-0.10	1.45	1.35	0.00	3	97	0	0.68	אנגלית	
1.29	-0.35	0.94	0.14	0	67	33	0.77	סכס	כמותיים
2.15	-1.21	0.94	0.33	0	47	53	0.88	בגרות	
0.07	0.87	0.94	0.01	0	98	2	0.52	מכפ"ל	
0.39	0.55	0.94	0.02	0	96	4	0.69	מילולי	
0.66	0.28	0.94	0.06	0	85	15	0.67	כמותי	
-0.74	1.68	0.94	0.00	1	99	0	0.51	אנגלית	
1.37	-0.22	1.15	0.16	0	66	34	0.88	סכס	כולם
2.24	-1.09	1.15	0.40	0	41	59	0.98	בגרות	
0.25	0.90	1.15	0.00	0	98	2	0.61	מכפ"ל	
0.48	0.67	1.15	0.01	0	96	4	0.75	מילולי	
0.81	0.34	1.15	0.05	0	87	13	0.80	כמותי	
-0.41	1.56	1.15	0.00	2	98	0	0.60	אנגלית	

^a במונחי סטיות תקן של הקריטריון; ערך חיובי מבטא הטיה לטובת נבחני ערבית וערך שלילי הטיה לרעתם.
^b ממוצע נבחני עברית פחות ממוצע נבחני ערבית

על פי המודל של קלירי כל החזאים מוטים לטובת נבחני ערבית. ממדי ניבוי היתר הם הגבוהים ביותר ביחס לבגרות (כסטיית תקן אחת). ניבוי היתר ביחס למכפ"ל הוא כ-0.6 סטיות תקן, והשילוב של שני החזאים מוביל לניבוי יתר של כ-0.9 סטיות תקן ביחס לציון הסכס. מבין מרכיבי מכפ"ל ממדי ניבוי היתר הם הקטנים ביותר ביחס לאנגלית (כ-0.6 סטיות תקן). ממדי ניבוי היתר ביחס לשני התחומים האחרים דומים זה לזה (כ-0.8 סטיות תקן). בהשוואה בין תחומי הלימוד רואים שממדי ניבוי היתר גבוהים יותר בתחומי הלימוד המילוליים. הדברים אמורים ביחס לכל החזאים.

גישת תנאי הגבול מספקת שני סוגי סטטיסטים: אחוז החוגים שבהם נמצאה הטיה (וכיוונה) וגודל ההטיה.

התמונה על אחוז החוגים שבהם נמצאה הטיה מראה שוב שמידת ההטיה לטובת נבחני ערבית גבוהה מאוד בבגרות (59% מן החוגים²⁶), ואפסית במכפ"ל (2% מן החוגים). מידת ההטיה לטובה של ציון הסכס היא באמצע (34% מן החוגים). אנגלית הוא החזאי היחיד שבו מתקיימת נטייה מזערית (2% מן החוגים) של הטיה לרעת נבחני ערבית. ביחס לשני תחומי מכפ"ל האחרים מוצאים שההטיה לטובה ניכרת יותר בתחום הכמותי (13% מן החוגים) מאשר בתחום המילולי (4% מן החוגים).

²⁵ כאמור, הערכים עבור תנאי הגבול והפרש d-ים מתוקנים לקיצוץ תחום. ערכים נצפים מוצגים בלוח 13 בנספח 2.
²⁶ במקרים שבהם המגמה אינה עקבית בין חוגים, מדווח בטקסט ההפרש בין אחוז החוגים עם הטיה בכיוון אחד לבין אחוז החוגים עם הטיה בכיוון שני (המספר הגדול פחות המספר הקטן).

בהשוואה בין תחומי לימוד התמונה אינה עקבית בין החזאים : בבגרות ההטיה לטובת נבחני ערבית בולטת יותר בתחומי הלימוד המילוליים, ואילו במכפ"ל ובמרכיביו הנטייה להטיה לטובתם חזקה יותר בתחומי הלימוד הכמותיים.

המדד של גודל ההטיה על פי גישת תנאי הגבול מראה הטיה לטובת נבחני ערבית או חוסר הטיה. ההטיה החזקה ביותר לטובת נבחני ערבית מתקיימת בבגרות (כ-0.4 סטיות תקן), מכפ"ל אינו מוטה, וציון הסכם מוטה לטובתם במידה מועטה (כ-0.2 סטיות תקן). מבין תחומי מכפ"ל, התחום הכמותי מוטה לטובת נבחני ערבית, אך במידה מזערית.

כצפוי, כמו שראינו ביחס לסטטיסטי הראשון של גישת תנאי הגבול – אחוז החוגים שבהם נמצאה הטיה – בהשוואה בין תחומי לימוד התמונה אינה עקבית בין החזאים : בבגרות ההטיה לטובת נבחני ערבית בולטת יותר בתחומי הלימוד המילוליים, ואילו במכפ"ל ובמרכיביו הנטייה להטיה לטובתם חזקה יותר בתחומי הלימוד הכמותיים.

על פי המדד של הפרש d-ים, כל החזאים, להוציא את אנגלית, מוטים לטובת נבחני ערבית. ההטיה לטובה של הבגרות מתבטאת בפער של כ-2.2 סטיות תקן בין ההפרש בין הקבוצות בממוצעי ציוניהן בקריטריון (כ-1.1 סטיות תקן לרעת נבחני ערבית) לבין ההפרש הני"ל בחזאי (כ-1.1 סטיות תקן לטובת נבחני ערבית). במכפ"ל הפער בין ה-d-ים הוא כ-0.3 סטיות תקן, ובציון הסכם הוא בין שני הערכים הללו (כ-1.4 סטיות תקן). מבין תחומי מכפ"ל מוצאים שאנגלית מוטה לרעת נבחני ערבית : ההפרש בין ה-d בקריטריון (כאמור, כ-1.2 סטיות תקן לרעת נבחני ערבית) קטן מההפרש בחזאי (כ-1.6 סטיות תקן לרעת נבחני ערבית). ביחס לשני התחומים האחרים מוצאים שההטיה לטובה ניכרת יותר בתחום הכמותי (פער של כ-0.8 סטיות תקן בין ה-d-ים) מאשר בתחום המילולי (פער של כ-0.5 סטיות תקן בין ה-d-ים).

בהשוואה בין תחומי הלימוד ניתן להתרשם שהנטייה להטיה לטובת נבחני ערבית בולטת יותר בתחומי הלימוד המילוליים.

התבוננות בו זמנית על שלושת המדדים לגודל ההטיה מובילה למסקנות הבאות :

1. כל החזאים על פי שלושת המדדים מוטים לטובת נבחני ערבית, להוציא את תחום האנגלית במכפ"ל שנמצא מוטה לרעת נבחני ערבית על פי המדד של הפרש d-ים.
2. בהתייחס לשלושת החזאים העיקריים של מערכת המיון מוצאים, על פי שלושת המדדים, שהבגרות היא החזאי שמוטה ביותר לטובת נבחני ערבית, מכפ"ל מוטה פחות, ומידת ההטיה של ציון הסכם היא באמצע.
3. בהתייחס לשלושת תחומי מכפ"ל מוצאים, על פי שלושת המדדים, שהתחום הכמותי הוא המוטה ביותר לטובת נבחני ערבית ואנגלית מוטה הכי פחות לטובתם.
4. כצפוי, בהשוואה בין שלושת המדדים ניתן לראות שתנאי הגבול מאתרים הטיה מזערית בהשוואה לשני המדדים האחרים. בהשוואה בין המדד של קלירי למדד של הפרש d-ים רואים שממדי ניבוי היתר על פי המדד של קלירי גדולים מממדי ניבוי היתר על פי המדד של הפרש d-ים

ביחס למכפ"ל ולשניים ממרכיביו (מילולי ואנגלית) וקטנים יותר ביחס לבגרות ולציון סכם. בנספח 3 מוצגים המתאמים בין ממדי ההטיה על פי שלושת המדדים.

מחקר 2 – מגדר

בלוח 7 מוצגים סטטיסטיים תיאוריים של משתני המחקר.

לוח 7

ממוצעים וסטיות תקן של החזאים והקריטריון לפי מגדר ותחומי לימוד

משתנה	גברים		נשים		תחומי לימוד
	ממוצע	סטיות תקן	ממוצע	סטיות תקן	
צש"א	82.9	8.3	83.2	8.1	מילוליים
סכם	56.1	7.6	56.4	6.9	
בגרות	95.3	7.7	98.4	6.8	
מכפ"ל	615.6	67.9	591.1	63.2	
מילולי	120.0	14.3	117.7	13.6	
כמותי	119.3	14.0	113.6	13.1	
אנגלית	122.8	17.1	117.3	16.2	
צש"א	79.2	10.4	78.6	10.2	כמותיים
סכם	62.8	5.1	63.5	4.9	
בגרות	102.1	6.1	104.6	5.4	
מכפ"ל	668.9	44.7	655.7	45.9	
מילולי	126.7	11.2	126.1	11.1	
כמותי	132.7	8.9	129.5	9.2	
אנגלית	129.6	14.2	125.8	13.9	
צש"א	81.0	9.5	80.8	9.2	כולם
סכם	59.5	6.4	60.0	6.0	
בגרות	98.7	7.0	101.6	6.2	
מכפ"ל	642.7	57.3	623.9	55.1	
מילולי	123.4	12.8	122.0	12.4	
כמותי	126.1	11.7	121.6	11.3	
אנגלית	126.2	15.7	121.6	15.1	

הממצאים המוצגים בלוח 7 מעידים שבציון שנה א' אין הבדלים משמעותיים בין הקבוצות. בבגרות יש יתרון לנשים (כ-0.4 סטיות תקן) ובמכפ"ל יש יתרון לגברים (כ-0.3 סטיות תקן). כתוצאה מכך, בציון הסכם דומים ממוצעי שתי הקבוצות. מבין תחומי מכפ"ל הפער בין הקבוצות (לטובת גברים) הוא הגדול ביותר בתחום הכמותי והקטן ביותר בתחום המילולי.

אין הבדלים משמעותיים בין תחומי הלימוד בעניין זה.

בלוח 8 מוצגים מקדמי המתאם בין החזאים לקריטריון.

לוח 8

מקדמי התוקף של החזאים לפי מגדר ותחומי לימוד²⁷

תחומי לימוד	חזאי	גברים	נשים
מילוליים	סכם	0.41	0.45
	בגרות	0.31	0.31
	מכפ"ל	0.40	0.47
	מילולי	0.35	0.41
	כמותי	0.32	0.37
	אנגלית	0.30	0.38
כמותיים	סכם	0.49	0.50
	בגרות	0.42	0.40
	מכפ"ל	0.42	0.46
	מילולי	0.31	0.35
	כמותי	0.39	0.43
	אנגלית	0.29	0.34
כולם	סכם	0.46	0.47
	בגרות	0.38	0.35
	מכפ"ל	0.41	0.46
	מילולי	0.33	0.38
	כמותי	0.36	0.39
	אנגלית	0.29	0.36

הממצאים המוצגים בלוח 8 מעידים שתוקף ציון הסכם דומה בשתי הקבוצות. תוקף הניבוי של מכפ"ל ומרכיביו גבוה יותר בקרב נשים. בבגרות הכיוון מתהפך והתוקף מעט יותר גבוה בקרב הגברים.

אין הבדלים משמעותיים בין תחומי הלימוד בעניין זה.

בלוח 9 מוצגים שלושת המדדים להטיה בבררה כלפי נשים.

²⁷ מתוקנים לקיצוץ תחום. מקדמי תוקף נציפים מוצגים בלוח 14 בנספח 2.

הפרש d-ים			תנאי גבול				קלירי ^a	חזאי	תחומי לימוד
d קריטי פחות חזאי ^a	d חזאי ^b	d קריטי ^b	גודל ההטיה ^a	% חוגים שבהם					
				הטיה לרעה	אין הטיה	הטיה לטובה			
0.01	-0.06	-0.05	0.00	12	76	12	0.00	סכס	מילוליים
0.38	-0.43	-0.05	0.04	2	72	26	0.03	בגרות	
-0.39	0.34	-0.05	-0.07	32	66	2	-0.05	מכפ"ל	
-0.20	0.15	-0.05	0.00	20	73	7	-0.02	מילולי	
-0.45	0.40	-0.05	-0.07	33	65	2	-0.04	כמותי	
-0.38	0.33	-0.05	-0.05	32	65	3	-0.04	אנגלית	
0.22	-0.31	-0.09	0.01	5	78	17	0.07	סכס	כמותיים
0.43	-0.52	-0.09	0.04	3	68	29	0.10	בגרות	
-0.16	0.07	-0.09	0.00	13	84	3	0.01	מכפ"ל	
0.01	-0.10	-0.09	0.00	7	87	6	0.04	מילולי	
-0.28	0.19	-0.09	0.00	18	80	2	-0.01	כמותי	
-0.26	0.17	-0.09	0.00	10	88	2	0.02	אנגלית	
0.12	-0.19	-0.07	0.00	8	77	15	0.04	סכס	כולם
0.41	-0.48	-0.07	0.04	2	70	28	0.07	בגרות	
-0.27	0.20	-0.07	0.00	22	75	3	-0.02	מכפ"ל	
-0.09	0.02	-0.07	0.00	13	81	6	0.01	מילולי	
-0.36	0.29	-0.07	-0.06	25	73	2	-0.02	כמותי	
-0.32	0.25	-0.07	0.00	20	77	3	-0.01	אנגלית	

^a במונחי סטיות תקן של הקריטריון; ערך חיובי מבטא הטיה לטובת נשים וערך שלילי הטיה לרעתן.
^b ממוצע גברים פחות ממוצע נשים

על פי המודל של קלירי, ביחס לכל החזאים, ממדי ההטיה הם מזעריים (פחות מ-0.1 סטיות תקן) והמערכת היא באופן כללי הוגנת.

גישת תנאי הגבול מספקת, כאמור, שני סוגי סטטיסטיים: אחוז החוגים שבהם נמצאה הטיה (וכיוונה) וגודל ההטיה.

התמונה על אחוז החוגים שבהם נמצאה הטיה מראה הטיה לטובת נשים על ידי הבגרות (26% מן החוגים²⁹), הטיה לרעתן על ידי מכפ"ל (19% מן החוגים) ונטייה קלה של הטיה לטובתן על ידי ציון הסכס (7% מן החוגים). מבין מרכיבי מכפ"ל מוצאים שההטיה לרעת נשים היא החזקה ביותר בתחום הכמותי (23% מן החוגים) והחלשה ביותר בתחום המילולי (7% מן החוגים).

בהשוואה בין תחומי לימוד התמונה עקבית מעבר לחזאים (אם כי ביחס לבגרות ההבדלים בין התחומים קטנים מאוד): ההטיה לרעה ניכרת יותר וההטיה לטובה ניכרת פחות בתחומי הלימוד המילוליים.

המדד של גודל ההטיה על פי גישת תנאי הגבול מראה, ביחס לכל החזאים, ממדי הטיה מזעריים (פחות מ-0.1 סטיות תקן) והמערכת היא באופן כללי הוגנת.

על פי המדד של הפרש d-ים, בגרות מוטה לטובת נשים ומכפ"ל מוטה לרעתן: בעוד שההפרש בין הקבוצות בממוצעי ציוניהן בקריטריון הוא מזערי, הפער ביניהן בבגרות הוא כ-0.5 סטיות תקן

²⁸ כאמור, הערכים עבור תנאי הגבול והפרש d-ים מתוקנים לקיצוץ תחום. ערכים נצפים מוצגים בלוח 15 בנספח 2.
²⁹ במקרים שבהם המגמה אינה עקבית בין חוגים, מדווח בטקסט ההפרש בין אחוז החוגים עם הטיה בכיוון אחד לבין אחוז החוגים עם הטיה בכיוון שני (המספר הגדול פחות המספר הקטן).

לטובת הנשים, ובמכפ"ל כ-0.2 סטיות תקן לרעת הנשים. הפער בין הקבוצות בציון הסכם הוא כ-0.2 סטיות תקן לטובת הנשים, לפיכך ציון הסכם מוטה מעט לטובתן. מבין תחומי מכפ"ל, התחום הכמותי ותחום האנגלית מוטים לרעת נשים (פער של כ-0.3 סטיות תקן לרעת נשים בכל אחד משני החזאים לעומת פער מזערי – לטובת נשים – בקריטריון).

בהשוואה בין תחומי הלימוד רואים גם במדד זה שההטיה לרעת נשים ניכרת יותר וההטיה לטובתן ניכרת פחות בתחומי הלימוד המילוליים.

התבוננות בו זמנית על שלושת המדדים לגודל ההטיה מובילה למסקנות הבאות:

1. בהתייחס לשלושת החזאים העיקריים של מערכת המיון מוצאים, על פי שלושת המדדים, שהבגרות מוטה לטובת נשים, מכפ"ל מוטה לרעתן, אך מידת ההטיה פחותה מזו של הבגרות, כך שציון הסכם מוטה מעט לטובתן.

2. בהתייחס לשלושת תחומי מכפ"ל מוצאים, על פי שלושת המדדים, שהתחום הכמותי הוא המוטה ביותר לרעת נשים, אנגלית מוטה אף הוא לרעתן, אם כי במידה פחותה, והתחום המילולי מוטה במידה מזערית.

3. בהשוואה בין שלושת המדדים ניתן להתרשם שהמדדים של קלירי ותנאי הגבול מאתרים הטיה מזערית בהשוואה למדד של הפרש d-ים. בנספח 3 מוצגים המתאמים בין ממדי ההטיה על פי שלושת המדדים.

סיכום ודיון

התבוננות בו זמנית בממצאי שני המחקרים, הגם שהם עוסקים, לכאורה, בנושאים שונים, מספקת קנה מידה לבחינת הערכים שהתקבלו בסטטיסטיים השונים, ומסייעת להתרשמות מכיוון הממצאים ומעוצמתם. הממצאים העיקריים (מעבר לתחומי לימוד) יסוכמו להלן בלוחות 10 (תוקף דיפרנציאלי) ו-11 (הטיה בבררה).

בלוח 10 מוצגת, עבור כל מחקר, הקבוצה עם התוקף הגבוה ביותר והפער בין התוקף בקרבה לבין התוקף בקבוצה המשלימה.

לוח 10

ממצאים מבדיקת התוקף הדיפרנציאלי בשני המחקרים

חזאי	התוקף גבוה יותר בקרב	הפער בתקפויות
מחקר 1 – שפת היבחנות		
סכס	עברית	0.14
בגרות	עברית	0.15
מכפ"ל	עברית	0.09
מילולי	עברית	0.13
כמותי	עברית	0.10
אנגלית	עברית	0.01
מחקר 2 – מגדר		
סכס	נשים	0.01
בגרות	גברים	0.03
מכפ"ל	נשים	0.05
מילולי	נשים	0.05
כמותי	נשים	0.03
אנגלית	נשים	0.07

בבדיקת תוקף דיפרנציאלי במחקר 1 נמצא שתוקף הניבוי של כל החזאים גבוה יותר בקרב נבחני עברית. הפערים בין הקבוצות בתקפויות הם ניכרים. במחקר 2 נמצא שההבדלים בתקפויות בין גברים לנשים הם קטנים ולא לגמרי עקביים בכיונם.

באף אחד מן המחקרים לא נמצאו הבדלים משמעותיים בין תחומי הלימוד בעניין זה.

לוח 11 מרוכזים הממצאים העיקריים מבדיקת ההטיה בבררה.

לוח 11

ממצאים מבדיקת הטיה בבררה כלפי קבוצת המוקד^a

חזאי	קלירי	תנאי גבול ^b	הפרש d-ים ^c
מחקר 1 (קבוצת המוקד : נבחני ערבית)			
סכס	0.88	0.16	1.37
בגרות	0.98	0.40	2.24
מכפ"ל	0.61	0.00	0.25
מילולי	0.75	0.01	0.48
כמותי	0.80	0.05	0.81
אנגלית	0.60	0.00	-0.41
מחקר 2 (קבוצת המוקד : נשים)			
סכס	0.04	0.00	0.12
בגרות	0.07	0.04	0.41
מכפ"ל	-0.02	0.00	-0.27
מילולי	0.01	0.00	-0.09
כמותי	-0.02	-0.06	-0.36
אנגלית	-0.01	0.00	-0.32

^a במונחי סטיות תקן של הקריטריון ; ערך חיובי מבטא הטיה לטובת קבוצת המוקד וערך שלילי הטיה לרעתה.

^b מוצג רק המדד לגודל ההטיה.

^c מוצג רק d קריטריון פחות d חזאי.

התמונה הכללית המתקבלת במחקר זה היא שהחזאי המרכזי של מערכת המיון, ציון הסכס, מוטה מאוד לטובת נבחני ערבית ומוטה מעט לטובת נשים. ניתן לסכס, לפיכך, וזאת בהתייחס לשני קבוצות המוקד שנבדקו במחקר זה, שתהליך המיון אינו פוגע (בהתייחס לנשים) ובמידה

רבה אף מעלה (בהתייחס לנבחני ערבית) את הנגישות של קבוצות אלה להשכלה הגבוהה. חשוב לחזור ולהדגיש שממצא זה התקבל על-פי שלושת המדדים להטיה בבררה, מה שמעיד על איתנותו.

בהתייחס לחזאים הנוספים של מערכת המיון התמונה היא כדלקמן :

בהתייחס לנבחני ערבית, הבגרות מוטה מאוד לטובתם ומכפ"ל אינו מוטה או מוטה פחות לטובתם (תלוי בהגדרה של הטיה בבררה), כאשר מבין מרכיבי מכפ"ל, התחום הכמותי הוא המוטה ביותר לטובת נבחני ערבית.

בהתייחס לנשים, הבגרות מוטה לטובתן ומכפ"ל מוטה לרעתן (אך מידת ההטיה שלו פחותה מזו של הבגרות), כאשר מבין מרכיבי מכפ"ל, התחום הכמותי הוא המוטה ביותר לרעתן.

התבוננות בו זמנית בממצאים בהתייחס לנבחני ערבית ולנשים מלמדת שהאפקטים המדוברים גדולים יותר כאשר קבוצת המוקד היא נבחני ערבית. ההבדל בין המחקרים ניכר בעיקר ביחס לבגרות ולציון הסכם³⁰. בדיקת האפקט המשולב של שני משתני הרקע (ראו נספח 4) מלמדת שהאפקט של מגדר דומה בשתי הקבוצות של שפת היבחנות והאפקט של שפת היבחנות דומה בשתי קבוצות המגדר.

ממצאי המחקר הנוכחי מתיישבים במידה רבה עם ממצאי המחקרים שבוצעו בעבר במאל"ו בנושאים אלו. כך, למשל, בהתייחס להוגנות על-פי שפת היבחנות, נמצא בעבר, בדומה למה שהתקבל במחקר הנוכחי, שתוקף החזאים גבוה יותר בקרב נבחני ערבית. בהיבט של הטיה בבררה, נמצא בעבר באופן עקבי שציון הסכם מוטה לטובת נבחני ערבית, כאשר הבגרות מוטה לטובתם במידה רבה מאוד ואילו מכפ"ל מוטה במידה מועטה, כשכיוון ההטיה הוא תלוי הגדרה. ממצאים אלה הם התקבלו במחקר הנוכחי (שבו, ראוי להדגיש, מכפ"ל לא נמצא מוטה לרעה לפי אף אחת מההגדרות). בהתייחס למרכיבי מכפ"ל, נמצא בעבר שכיוון ההטיה הוא תלוי הגדרה, אך מבין השלושה, התחום הכמותי מוטה יותר לטובת נבחני ערבית. ממצא זה התקבל במחקר הנוכחי. בהתייחס להוגנות על-פי מגדר, נמצאו בעבר הבדלים קטנים בתקפיות בין גברים לנשים, שכיוונם לא עקבי. ממצאים דומים התקבלו במחקר הנוכחי. בהיבט של הטיה בבררה נמצא בעבר שציון הסכם אינו מוטה, כתוצאה מכך שבגרות ומכפ"ל פועלים בכיוונים מנוגדים : הבגרות מוטה לטובת נשים ומכפ"ל מוטה לרעתן. ממצאים דומים התקבלו במחקר הנוכחי (עם נטייה להטיה לטובה גם על ידי ציון הסכם, כתוצאה מכך שמידת ההטיה במכפ"ל פחותה מזו של הבגרות). בהתייחס למרכיבי מכפ"ל, נמצא בעבר שהתחום הכמותי הוא המוטה ביותר לרעת נשים. ממצא זה התקבל במחקר הנוכחי.

ממצאי המחקר הנוכחי דומים ברוחם גם לממצאי מחקרים על הוגנות בבררה שנעשו ביחס ל-Matter et al., 2008) SAT. בפרט, נמצא במחקרים אלה שתוקף חזאי מערכת המיון נמוך יותר בקרב קבוצות אתניות וגבוה יותר בקרב נשים, ושיש הטיה בבררה לטובת קבוצות מיעוט אתניות ולרעת נשים.

³⁰ במדד של קלירי מתקיים הבדל כזה גם ביחס למכפ"ל ולמרכיביו. יש לזכור שממדי ההטיה על פי מדד זה, כפי שיושם כאן, תלויים בגודלה היחסי של קבוצת המוקד : כשיתר הדברים (ובפרט, המרחק בין קו הרגרסיה של קבוצת המוקד לקו הרגרסיה של קבוצת ההתייחסות) שווים, ככל שקבוצת המוקד גדולה יותר, קו הרגרסיה המשותף יהיה קרוב יותר לקו הרגרסיה (הנפרד) שלה.

מה מקורם של הממצאים שהתקבלו בשני המחקרים? בהתייחס לשפת היבחנות, חשוב לזכור שבעוד שכלי המיון בודקים את היכולת הלימודית בשפת האם של נבחני ערבית, ההישגים בלימודים עצמם נמדדים בעברית, שבה רמת השליטה שלהם נמוכה יותר. אין זה מפתיע, לפיכך, למצוא שכלי המיון מוטים לטובת נבחני ערבית. תמיכה להסבר זה ניתן לקבל מן הממצא, שהתקבל במחקר זה, שבתחומי הלימוד המילוליים מידת ההטיה לטובת נבחני ערבית גדולה יותר מאשר בתחומי לימוד כמותיים. במילים אחרות, ככל שהקריטריון טעון יותר בשפה, ההטיה של כלי המיון לטובה חזקה יותר. צוויק (Zwick, 2002; Zwick & Himelfarb, 2011) הציעה הסברים אחדים לממצא הרווח של הטיה לטובת קבוצות מיעוט אתניות בארה"ב, שעשויים להיות תקפים גם לנבחני ערבית. כך, למשל, מוצגת האפשרות שהפוטנציאל הלימודי של סטודנטים מקבוצות מיעוט נמדד כהלכה על ידי כלי המיון, אך הם לא מצליחים לממש אותו בלימודים עצמם. ה"ת-ביצוע" הזה עשוי לנבוע מגזענות ישירה או מאווירת קמפוס שאינה מסבירת פנים לקבוצות מיעוט, או שהוא קשור לחשיפה רבה יותר של חברי קבוצות מיעוט לנסיבות חיים מכבידות, כולל קשיים כלכליים, שעשויות לחבל בתפקוד בלימודים. השערה נוספת שהועלתה היא שחרדות, שאיפות נמוכות ועמדות שליליות עשויות לפגוע בהצלחה האקדמית של חברי קבוצות מיעוט (Bowen & Bok, 1998).

בהתייחס למגדר, חשוב לחזור ולהדגיש שהאפקטים שנמצאו הם קטנים מאוד. עם זאת, מגמת הממצאים שהתקבלו ביחס לחזאים השונים מתיישבת עם הידע הקיים אודות הבדלי מגדר במדדי יכולת לימודית. בסקירתם המקיפה על הבדלי ביצוע בין גברים לנשים, דיווחו ווילינגהאם וקול (Willingham & Cole, 1997) על הישגים ממוצעים גבוהים יותר של נשים בציוני בית ספר ועל הישגים ממוצעים גבוהים יותר של גברים במבחנים מתוקננים, כאשר בתוך המבחנים המתוקננים, היתרון של הגברים ניכר יותר בחלקים הכמותיים של המבחן. בין היתר, הוצעו הסברים סטטיסטיים להבדלי הביצוע במבחנים מתוקננים (שונות גדולה יותר בציונים באוכלוסיית הגברים, מה שמביא לממצא של ממוצע גבוה יותר לגברים באוכלוסיות שמורכבות מבעלי היכולת הגבוהה). לענייננו, חשוב להתבונן בהבדלים אלה לנוכח אופיו של הקריטריון (ממוצע ציוני שנה א' בלימודי התואר הראשון). קריטריון זה שונה באופיו מן המדידה המתקבלת במבחנים מתוקננים (בהיותו מבוסס יותר על מבחנים פתוחים ולא על מבחני בררה). עם זאת, הוא שונה גם מציוני בית ספר בכך שהוא מתבסס בעיקר על מבחנים (או עבודות) ואינו כולל מרכיבי הערכה סובייקטיביים, כדוגמת מוטיבציה, עניין והשתתפות פעילה. ממצאי ההטיה בבררה (של מכפ"ל ובגרות) ביחס לקריטריון זה משקפים את היותו תוצר של מדידה משולבת, הן זו המאפיינת ציוני בית ספר והן זו המאפיינת ציונים במבחנים מתוקננים. בהשוואה בין תחומי לימוד נמצא שההטיה לרעת נשים ניכרת יותר וההטיה לטובת נשים ניכרת פחות בתחומים המילוליים. ממצא זה קשור לדפוס הביצוע שתוארו לעיל בחלקים המילוליים והכמותיים במבחנים המתוקננים: היכולת הגבוהה, יחסית, של נשים בחלקים המילוליים אינה רלוונטית לקריטריון בתחומי לימוד כמותיים (מה שגורר הטיה לטובת נשים בתחומי לימוד כמותיים), והיכולת הגבוהה של גברים בחלקים הכמותיים אינה רלוונטית לקריטריון בתחומי לימוד מילוליים (מה שגורר הטיה לטובת גברים ולרעת נשים בתחומי לימוד מילוליים).

לבסוף, בהתייחס לנושא המתודולוגי של עבודה זו – השוואה בין שלושה מדדים להטיה בבררה – התמונה המתקבלת היא של מתאמים חיוביים לא מבוטלים בין ממדי ההטיה על פי שלושת המדדים (ראו נספח 3), ובפרט עקביות בין המדדים באפיון החזאים השונים מבחינת כיוון ההטיה ובדירוגם על פי ממדי ההטיה שהם יוצרים. כצפוי, המדד של תנאי הגבול נוטה הכי פחות לאתר הטיה. זהו מדד שמרני, שמאפשר רק למצבים קיצוניים של הטיה להיות מזוהים ככאלה. חשוב להדגיש, עם זאת, שבהשוואה למודל הרגרסיה של קלירי, מופעלת על ידי גישת תנאי הגבול "צנזורה" על מסקנות של הטיה רק במצבים של הטיה לטובת הקבוצה החלשה (שהם המצבים שנוחים לממסד הבוחן או הממין). כל המצבים שבהם, על פי המודל של קלירי, יש הטיה לרעת הקבוצה החלשה, יזוהו ככאלה גם בגישת תנאי הגבול. לפיכך, יחסית לגישה הרווחת לאיתור הטיה בבררה, מייצגת גישת תנאי הגבול עמדה מחמירה כלפי הממסד הבוחן או הממין. היחס בין ממדי ההטיה על פי שני המדדים האחרים – המדד של קלירי והמדד של הפרש d-ים – אינו עקבי והוא תלוי בפער בין הקבוצות בחזאי: במצב (השכיח) שבו קבוצת המוקד היא הקבוצה החלשה יותר (נמוכה יותר בחזאי), המדד של קלירי נוטה להראות יותר הטיה לטובתה (או פחות הטיה לרעתה) מאשר המדד של הפרש d-ים, ולהיפך במצב שבו קבוצת המוקד היא הקבוצה החזקה יותר.

חשוב לחזור ולהדגיש שההבדלים בין המדדים קשורים לא רק להבדלים בהגדרות, אלא גם להבדלים באופרציונליזציה שלהן במחקר הנוכחי. כך, למשל, המדד של קלירי לא תוקן לקיצוץ תחום והמדדים של תנאי הגבול ושל הפרש d-ים כן (הסיבה לאי תיקון הראשון הוצגה במבוא; תוצאות לא מתוקנות של השניים האחרים מוצגות בנספח 2). ואמנם, הבאת ההגדרות למכנה משותף בנושא זה הגדילה את המתאמים בין המדדים (ראו נספח 3).

הבדל טכני נוסף בין המדדים – הפעם בין המדד של תנאי הגבול לשני המדדים האחרים – התקיים ביישום תנאי הגבול במחקרים קודמים במאלי"ו, וטופל במחקר הנוכחי. באופן מפורט, בצבירת המידע הנומינלי שהתקבל בעבר בגישת תנאי הגבול ("הטיה לטובה", "הטיה לרעה" או "לא נמצאה הטיה") ניתן משקל זהה לכל חוג ולא הובאו בחשבון הבדלים בגדלי החוגים. במדד החדש לגודל ההטיה בגישת תנאי הגבול משוקלל הממצא החוגי במספר הסטודנטים בחוג, כמו בגישות האחרות.

כיוון נוסף, שראוי לבדוק בעתיד במטרה לצמצם הבדלים באופרציונליזציה של המדדים, הוא הטיפול בהבדלי שיפועים בין קווי הרגרסיה של קבוצת המוקד לקבוצת ההתייחסות ביישום הגישה של קלירי וביישום תנאי הגבול. באופן מפורט, בעוד שהמדד של קלירי, ביישום הנוכחי, מתעלם מהבדלים כאלה, המדד של גישת תנאי הגבול מביא בחשבון הבדלים אלו (ראו הערה 22), דבר שמקטין את הנטייה להסיק על הטיה ובכך מחזק את השמרנות של מדד זה בהשוואה למדד של קלירי.

כאמור, לצרכים מתודולוגיים יש לשאוף לעקביות רבה ככל האפשר ביישום ההגדרות השונות להטיה בבררה. עם זאת, המסקנות המהותיות מתבררות גם בתנאים הנוכחיים: מערכת המיון מוטה לטובת נבחני ערבית והוגנת לשני המגדרים.

מקורות

- ברון, ה' וגפני, נ' (1989). **הוגנות השימוש בבחינה הפסיכומטרית עבור הנבחנים בעברית ובערבית** (דוח מס' 93). ירושלים: מרכז ארצי לבחינות ולהערכה.
- ברונר, ש' (2004). **השלכות שילוב מכפ"ל או מצרף במערכת המיון להשכלה הגבוהה על תוקף הניבוי וההוגנות של המערכת** (דוח טכני מס' 152). ירושלים: מרכז ארצי לבחינות ולהערכה.
- ברונר, ש', אללוף, א' ואורן, כ' (1996). **הוגנות מערכת המיון לאוניברסיטאות עבור נבחני מכפ"ל בערבית – השפעת כיוול המרכיב המילולי** (דוח מס' 223). ירושלים: מרכז ארצי לבחינות ולהערכה.
- גמליאל, א' וקאהן, ס' (2004). **חוסר הוגנות בקבלה לאוניברסיטאות: הגרסה המודרנית של משל "כבשת הרש"'**. מגמות, מג, 445-433.
- גפני, נ' ובלר, מ' (1989). **בדיקת הוגנות השימוש בבחינה הפסיכומטרית עבור שני המינים** (דוח מס' 95). ירושלים: מרכז ארצי לבחינות ולהערכה.
- נבון, ד' וכהן, י' (2009). **אפליה בבחינות הכניסה לאוניברסיטאות או ארטיפקטים סטטיסטיים: תגובה לגמליאל וקאהן** (2004). **מגמות, מו, 418-398**.
- קנת-כהן, ת' (2001). **ניבוי דיפרנציאלי ותוקף דיפרנציאלי של מערכת המיון לאוניברסיטאות על פי מיצב חברתי-כלכלי של המועמדים** (דוח מס' 285א). ירושלים: מרכז ארצי לבחינות ולהערכה.
- American Educational Research Association, American Psychological Association & National Council on Measurement in Education (2011, Draft for revision). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.
- Azen, R., Bronner, S., & Gafni, N. (1999). *Examination of gender bias in admission to universities in Israel* (Report No. 255). Jerusalem: National Institute for Testing and Evaluation.
- Azen, R., Bronner, S., & Gafni, N. (2002). Examination of gender bias in university admissions. *Applied Measurement in Education, 15*, 75-94.
- Birnbaum, M. H. (1979). Procedures for detection and correction of salary inequity. In T. R. Pezzullo & B. F. Birthingam (Eds.), *Salary Equity* (pp. 121-144). Lexington, MA: Lexington Books.
- Birnbaum, M. H. (1981). Reply to McLaughlin: Proper path models for theoretical partialling. *American Psychologist, 36*, 1193-1195.
- Bowen, W. G., & Bok, D. (1998). *The shape of the river: Long-term consequences of considering race in college and university admissions*. Princeton, NJ: Princeton University Press.
- Cahan, S., & Gamliel, E. (2006). Definition and measurement of selection bias: From constant ratio to constant difference. *Journal of Educational Measurement, 43*, 131-144.
- Camilli, J. (2006). Test fairness. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 221-256). Westport: American Council on Education & Praeger.
- Cleary, T. A. (1968). Test bias: Prediction of grades of Negro and white students in integrated colleges. *Journal of Educational Measurement, 5*, 115-124.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Cole, N. S. (1973). Bias in selection. *Journal of Educational Measurement, 10*, 237-255.
- Cole, N. S., & Moss, P. A. (1989). Bias in test use. In R. L. Linn (Ed.), *Educational*

- measurement* (3rd ed., pp. 201-220). New York: American Council on Education & Macmillan.
- Darlington, R. B. (1971). Another look at "culture fairness". *Journal of Educational Measurement*, 8, 71-81.
- Gafni, N., Beller, M., & Bronner, S. (2000). *A cross-cultural perspective on gender differences in higher education: Admissions and scholastic achievement* (Report No. 274). Jerusalem: National Institute for Testing and Evaluation.
- Gafni, N., & Bronner, S. (1998). *An examination of criterion-related bias in the testing of Hebrew- and Russian-speaking examinees in Israel* (Report No. 244). Jerusalem: National Institute for Testing and Evaluation.
- Gulliksen, H. (1950). *Theory of mental tests*. New York: John Wiley & Sons. {Reprinted in 1987. Hillsdale, NJ: Erlbaum.}
- Hunter, J. E., & Schmidt, F. L. (1976). Critical analysis of the statistical and ethical implications of various definitions of test bias. *Psychological Bulletin*, 83, 1053-1071.
- Kane, M. T., & Mroch, A. A. (2010). Modeling group differences in OLS and orthogonal regression: Implications for differential validity studies. *Applied Measurement in Education*, 23, 215-241.
- Linn, R. L. (1973). Fair test use in selection. *Review of Educational Research*, 43, 139-161.
- Linn, R. L. (1983). Pearson selection formulas: Implications for studies of predictive bias and estimates of educational effects in selected samples. *Journal of Educational Measurement*, 20, 1-15.
- Linn, R. L. (1984). Selection bias: Multiple meanings. *Journal of Educational Measurement*, 21, 33-47.
- Mattern, K. D., Patterson, B. F., Shaw, E. J., Kobrin, J. L., & Barbuti, S. M. (2008). *Differential validity and prediction of the SAT* (College Board Research Report No. 2008-4). New York: The College Board.
- Peterson, N. S., & Novick, M. R. (1976). An evaluation of some models for culture-fair selection. *Journal of Educational Measurement*, 13, 3-29.
- Thorndike, R. L. (1971). Concepts of culture-fairness. *Journal of Educational Measurement*, 8, 63-70.
- Turvall, E., Bronner, S., Kennet-Cohen, T., & Oren, C. (2008). *Fairness in the higher education admissions procedure: The Psychometric Entrance Test in Arabic* (Report No. 349). Jerusalem: National Institute for Testing and Evaluation.
- Willingham, W. W., & Cole, N. S. (1997). *Gender and fair assessment*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Young, J. W. (2001). *Differential validity, differential prediction, and college admission testing: A comprehensive review and analysis* (College Board Research Report No. 2001-6). New York: The College Board.
- Zwick, R. (2002). *Fair game? The use of standardized admissions tests in higher education*. New York, NY: RoutledgeFalmer.
- Zwick, R., & Himelfarb, I. (2011). The effect of high school socioeconomic status on the predictive validity of SAT scores and high school grade-point average. *Journal of Educational Measurement*, 48, 101-121.

נספח 1

חישוב המדד לגודל ההטיה על פי תנאי הגבול

ממדי ניבוי היתר או החסר הוגדרו כהפרש בין שני ערכים של הקריטריון: ערך הקריטריון כאשר הציון בחזאי שווה לציון הממוצע בחוג על פי משוואת הרגרסיה של קבוצת ההתייחסות וערך הקריטריון (במצב הנ"ל) על פי משוואת הרגרסיה של קבוצת המוקד. חשוב להבהיר שתי נקודות:

1. כאמור, ביישום תנאי הגבול, לא חושבו משוואות רגרסיה נפרדות לקבוצת ההתייחסות ולקבוצת המוקד. במקום זה, חושבה משוואת רגרסיה מרובה. המקדמים של משוואה זו מאפשרים לחשב את ההפרש הנ"ל, כפי שיפורט להלן.

2. כאמור, משוואת רגרסיה (מרובה) אחת הייתה משוואת הרגרסיה ה"רגילה" (לניבוי הקריטריון), ומשוואת רגרסיה (מרובה) שנייה הייתה משוואת הרגרסיה ה"הפוכה" (לניבוי החזאי). המשוואה שעל פיה חושב גודל ההטיה הייתה משוואת הרגרסיה שממנה התקבלה המסקנה על הטיה. בכל מקרה, ההפרש המחושב הוא בין שני ערכים של הקריטריון, במילים אחרות, מה שחושב הוא המרחק האנכי בין קווי הרגרסיה (של קבוצת ההתייחסות ושל קבוצת המוקד) שנגזרים ממשוואת הרגרסיה המרובה הקובעת, בין אם הרגרסיה המרובה הקובעת הייתה הרגרסיה ה"רגילה" ובין אם היא הייתה הרגרסיה ה"הפוכה".

להלן מתואר פיתוח הנוסחה שעל פיה חושב המדד להטיה בשני המקרים: כאשר הרגרסיה הקובעת הייתה הרגרסיה ה"רגילה" וכאשר היא הייתה הרגרסיה ה"הפוכה".

בשני המקרים נגדיר:

$G = 1$ = משתנה של שיוך קבוצתי: אם זו קבוצת המוקד $G = 0$ ואם זו קבוצת ההתייחסות $G = 1$.

X = החזאי

Y = הקריטריון

כאשר משוואת הרגרסיה הקובעת היא הרגרסיה ה"רגילה" (Y מנובא על ידי X) –

משוואת הרגרסיה המרובה היא $Y = a_0G + a_1X + a_2GX$ (עד a_2 הם מקדמי הרגרסיה)

קו הרגרסיה של קבוצת המוקד שנגזר ממשוואת הרגרסיה המרובה: $Y = a_1X$

קו הרגרסיה של קבוצת ההתייחסות שנגזר ממשוואת הרגרסיה המרובה: $Y = a_0 + (a_1 + a_2)X$

לכן כאשר $X = 0$ (הציון הממוצע בחוג, מעבר לשתי הקבוצות) ההפרש בין הקווים הוא a_0 וזה

המדד לגודל ההטיה.

כאשר משוואת הרגרסיה הקובעת היא הרגרסיה ה"הפוכה" (X מנובא על ידי Y) –

משוואת הרגרסיה המרובה היא $X = b_0G + b_1Y + b_2GY$ (עד b_2 הם מקדמי הרגרסיה)

קו הרגרסיה של קבוצת המוקד שנגזר ממשוואת הרגרסיה המרובה: $X = b_1Y$

קו הרגרסיה של קבוצת ההתייחסות שנגזר ממשוואת הרגרסיה המרובה: $X = b_0 + (b_1 + b_2)Y$

באופן אנלוגי למה שהוצג קודם ביחס לרגרסיה ה"רגילה", כאשר $Y = 0$ ההפרש בין הקווים הוא b_0 , אבל זהו ההפרש האנפקי בין הקווים. אנחנו מבקשים להסתכל על התמונה מן הזווית של ניבוי יתר או חסר של הקריטריון, ולכן מתבוננים גם כאן במרחק האנכי בין הקווים (כאשר $X = 0$).

עבור קבוצת המוקד, כאשר $Y = 0 : X = 0$

עבור קבוצת ההתייחסות, כאשר $X = 0$ מתקיים: $0 = b_0 + (b_1 + b_2)Y$

כלומר $Y = -b_0 / (b_1 + b_2)$.

לכן כאשר $X = 0$ ההפרש בין הקווים הוא $-b_0 / (b_1 + b_2)$ וזה המדד לגודל ההטיה.

נספח 2

סטטיסטיים נצפים

לוח 12

מקדמי תוקף נצפים של החזאים לפי שפת היבחנות ותחומי לימוד

ערבית	עברית	חזאי	תחומי לימוד
0.26	0.39	סכם	מילוליים
0.20	0.34	בגרות	
0.21	0.28	מכפ"ל	
0.15	0.25	מילולי	
0.16	0.22	כמותי	
0.15	0.16	אנגלית	כמותיים
0.21	0.33	סכם	
0.19	0.29	בגרות	
0.09	0.17	מכפ"ל	
-0.02	0.08	מילולי	
0.14	0.22	כמותי	כולם
0.10	0.05	אנגלית	
0.24	0.36	סכם	
0.20	0.31	בגרות	
0.16	0.22	מכפ"ל	
0.07	0.16	מילולי	כולם
0.15	0.22	כמותי	
0.13	0.11	אנגלית	

לוח 13

שני מדדים להטיה בבררה כלפי נבחני ערבית, מחושבים ללא תיקון לקיצוץ תחום

הפרש d-ים			תנאי גבול			חזאי	תחומי לימוד	
d קריטי פחות	d חזאי ^b	d קריטי ^b	גודל ההטיה ^a	% חוגים שבהם				
d חזאי ^a				הטיה לרעה	הטיה הטובה			
1.47	0.03	1.50	0.19	0	64	36	מילוליים	
2.56	-1.06	1.50	0.45	0	32	68		בגרות
0.28	1.22	1.50	0.00	0	99	1		מכפ"ל
0.54	0.96	1.50	0.01	1	95	4		מילולי
0.98	0.52	1.50	0.05	1	87	12		כמותי
-0.12	1.61	1.50	0.00	3	97	0	אנגלית	כמותיים
1.22	-0.15	1.07	0.18	0	58	42	סכם	
2.41	-1.34	1.07	0.34	0	33	67	בגרות	
-0.25	1.31	1.07	0.01	0	98	2	מכפ"ל	
0.30	0.77	1.07	0.02	0	95	5	מילולי	
0.65	0.42	1.07	0.07	0	82	18	כמותי	כולם
-0.81	1.88	1.07	0.00	1	99	0	אנגלית	
1.35	-0.06	1.28	0.18	0	61	39	סכם	
2.48	-1.20	1.28	0.39	0	32	68	בגרות	
0.02	1.27	1.28	0.01	0	98	2	מכפ"ל	
0.42	0.86	1.28	0.02	1	95	4	מילולי	כולם
0.82	0.47	1.28	0.06	0	84	16	כמותי	
-0.46	1.75	1.28	0.00	2	98	0	אנגלית	

^a במונחי סטיות תקן של הקריטריון; ערך חיובי מבטא הטיה לטובת נבחני ערבית וערך שלילי הטיה לרעתם.
^b ממוצע נבחני עברית פחות ממוצע נבחני ערבית

לוח 14

מקדמי תוקף נצפים של החזאים לפי מגדר ותחומי לימוד

תחומי לימוד	חזאי	גברים	נשים
מילוליים	סכם	0.36	0.36
	בגרות	0.24	0.19
	מכפ"ל	0.35	0.38
	מילולי	0.31	0.32
	כמותי	0.26	0.28
	אנגלית	0.27	0.31
כמותיים	סכם	0.32	0.33
	בגרות	0.22	0.20
	מכפ"ל	0.22	0.27
	מילולי	0.12	0.14
	כמותי	0.25	0.28
	אנגלית	0.14	0.18
כולם	סכם	0.33	0.35
	בגרות	0.23	0.19
	מכפ"ל	0.27	0.34
	מילולי	0.19	0.25
	כמותי	0.25	0.28
	אנגלית	0.19	0.25

לוח 15

שני מדדים להטיה בבררה כלפי נשים, מחושבים ללא תיקון לקיצוץ תחום

תחומי לימוד	חזאי	תנאי גבול			הפרש d-ים			
		הטיה לטובה	אין הטיה	הטיה לרעה	גודל ההטיה ^a	d קריטי ^b	d חזאי ^b	d קריטי פחות חזאי ^a
						% חוגים שבהם		
מילוליים	סכם	13	75	12	0.00	-0.05	0.00	
	בגרות	28	70	2	0.04	-0.47	0.42	
	מכפ"ל	3	63	34	-0.08	0.41	-0.46	
	מילולי	7	70	23	0.00	0.17	-0.22	
	כמותי	2	63	35	-0.07	0.46	-0.50	
	אנגלית	3	64	33	-0.06	0.35	-0.40	
כמותיים	סכם	22	73	5	0.02	-0.15	0.18	
	בגרות	33	63	4	0.06	-0.46	0.49	
	מכפ"ל	4	80	16	0.00	0.30	-0.27	
	מילולי	8	83	9	0.00	0.05	-0.02	
	כמותי	3	75	22	0.00	0.37	-0.34	
	אנגלית	3	84	13	0.00	0.30	-0.27	
כולם	סכם	18	74	8	0.01	-0.10	0.09	
	בגרות	31	66	3	0.05	-0.47	0.46	
	מכפ"ל	3	72	25	-0.05	0.36	-0.36	
	מילולי	7	77	16	0.00	0.11	-0.12	
	כמותי	2	69	29	-0.06	0.41	-0.42	
	אנגלית	3	74	23	0.00	0.32	-0.33	

^a במונחי סטיות תקן של הקריטריון; ערך חיובי מבטא הטיה לטובת נשים וערך שלילי הטיה לרעתן.
^b ממוצע גברים פחות ממוצע נשים

נספח 3

מתאמים בין ממדי ההטיה על פי שלושת המדדים

המתאמים בין ממדי ההטיה על פי שלושת המדדים מוצגים בלוחות 16 (עבור מחקר 1) ו-18 (עבור מחקר 2). שלוש עמודות המספרים הימניות מתייחסות למדדים כפי שתוארו בדוח. מאחר שהמדד של קלירי אינו מתוקן לקיצוץ תחום, חושבו המתאמים גם במצב שבו שלושת המדדים אינם מתוקנים לקיצוץ תחום. הערכים שהתקבלו מוצגים בשלוש העמודות השמאליות. בנוסף, מאחר שאחוז החוגים שבהם לא נמצאה הטיה על פי תנאי הגבול הוא גבוה מאוד, היו חוגים רבים שנקבע להם ערך זהה של 0 במדד החדש לגודל ההטיה על פי תנאי הגבול. כדי להעריך את עוצמת הקשר במקרים שבהם נמצאה הטיה על פי תנאי הגבול, מוצגים בלוח 17 (עבור מחקר 1) ובלוח 19 (עבור מחקר 2) המתאמים בין שלושת מדדי ההטיה רק בחוגים אלה.

לוח 16

מתאמי פירסון בין שלושת מדדי ההטיה במחקר 1

כל המדדים מחושבים ללא תיקון לקיצוץ תחום			תיקון לקיצוץ תחום כפי שהוצג בדוח ^a			חזאי
תנאי גבול & הפרש d-ים	קלירי & הפרש d-ים	קלירי & תנאי גבול	תנאי גבול & הפרש d-ים	קלירי & הפרש d-ים	קלירי & תנאי גבול	
0.55	0.71	0.15	0.51	0.48	0.12	סכס
0.41	0.55	0.02	0.35	0.31	-0.12	בגרות
0.22	0.65	0.03	0.23	0.62	0.06	מכפ"ל
0.33	0.63	0.08	0.31	0.55	0.08	מילולי
0.39	0.75	0.09	0.35	0.66	0.08	כמותי
0.19	0.59	0.20	0.19	0.56	0.19	אנגלית

^a זכור, המדד של קלירי לא מתוקן לקיצוץ תחום והמדדים של תנאי הגבול והפרש d-ים מתוקנים לקיצוץ תחום.

לוח 17

מתאמי פירסון בין שלושת מדדי ההטיה ללא חוגים שבהם לא נמצאה הטיה על פי תנאי הגבול

במחקר 1

כל המדדים מחושבים ללא תיקון לקיצוץ תחום				תיקון לקיצוץ תחום כפי שהוצג בדוח ^a				חזאי
תנאי גבול & הפרש d-ים	קלירי & הפרש d-ים	קלירי & תנאי גבול	מספר חוגים	תנאי גבול & הפרש d-ים	קלירי & הפרש d-ים	קלירי & תנאי גבול	מספר חוגים	
0.81	0.73	0.51	148	0.70	0.35	0.39	126	סכס
0.63	0.55	0.26	255	0.55	0.25	0.05	221	בגרות
0.82	0.85	0.55	7	0.89	0.73	0.76	7	מכפ"ל
0.92	0.82	0.76	19	0.85	0.62	0.76	16	מילולי
0.77	0.74	0.46	60	0.73	0.54	0.42	50	כמותי
0.82	-0.11	0.14	8	0.87	-0.07	0.03	7	אנגלית

^a זכור, המדד של קלירי לא מתוקן לקיצוץ תחום והמדדים של תנאי הגבול והפרש d-ים מתוקנים לקיצוץ תחום.

התמונה המתקבלת מלוח 16 היא של מתאמים גבוהים למדי בין המדד של קלירי למדד של הפרש d-ים (בממוצע מעבר לחזאים : 0.53). המתאמים בין המדד של תנאי הגבול למדד של הפרש d-ים נמוכים מהם (בממוצע : 0.32). המתאמים בין המדד של תנאי הגבול למדד של קלירי אפסיים. כאשר כל המדדים אינם מתוקנים לקיצוץ תחום, עולים בדרך כלל המתאמים ביניהם, אך התמונה הכללית נשארת בעינה : מתאמים גבוהים בין המדד של קלירי למדד של הפרש d-ים,

מתאמים בינוניים בין המדד של תנאי הגבול למדד של הפרש d-ים ומתאמים נמוכים בין המדד של תנאי הגבול למדד של קלירי. כאשר המתאמים בין המדדים (כולם ללא תיקון לקיצוץ תחום) מחושבים רק בחוגים שבהם נמצאה הטיה על פי תנאי הגבול (שלוש העמודות השמאליות בלוח 17), מוצאים מתאמים גבוהים בהרבה³¹ מן המתאמים שהתקבלו מעבר לכל החוגים (שלוש העמודות השמאליות בלוח 16).

לוח 18

מתאמי פירסון בין שלושת מדדי ההטיה במחקר 2

כל המדדים מחושבים ללא תיקון לקיצוץ תחום			תיקון לקיצוץ תחום כפי שהוצג בדוח ^a			חזאי
תנאי גבול & הפרש d-ים	קלירי & הפרש d-ים	קלירי & תנאי גבול	תנאי גבול & הפרש d-ים	קלירי & הפרש d-ים	קלירי & תנאי גבול	
0.71	0.75	0.56	0.60	0.60	0.50	סכס
0.63	0.66	0.60	0.52	0.56	0.54	בגרות
0.65	0.69	0.51	0.58	0.62	0.47	מכפ"ל
0.64	0.72	0.51	0.61	0.64	0.46	מילולי
0.67	0.66	0.52	0.60	0.63	0.44	כמותי
0.62	0.66	0.49	0.59	0.63	0.45	אנגלית

^a כזכור, המדד של קלירי לא מתוקן לקיצוץ תחום והמדדים של תנאי הגבול והפרש d-ים מתוקנים לקיצוץ תחום.

לוח 19

מתאמי פירסון בין שלושת מדדי ההטיה ללא חוגים שבהם לא נמצאה הטיה על פי תנאי הגבול

במחקר 2

כל המדדים מחושבים ללא תיקון לקיצוץ תחום				תיקון לקיצוץ תחום כפי שהוצג בדוח ^a				חזאי
תנאי גבול & הפרש d-ים	קלירי & הפרש d-ים	קלירי & תנאי גבול	מספר חוגים	תנאי גבול & הפרש d-ים	קלירי & הפרש d-ים	קלירי & תנאי גבול	מספר חוגים	
0.95	0.83	0.84	219	0.94	0.81	0.84	189	סכס
0.91	0.67	0.81	282	0.87	0.64	0.78	248	בגרות
0.91	0.76	0.85	235	0.91	0.73	0.84	206	מכפ"ל
0.95	0.85	0.89	190	0.91	0.78	0.89	160	מילולי
0.87	0.64	0.73	256	0.82	0.66	0.69	226	כמותי
0.91	0.76	0.80	215	0.90	0.73	0.81	188	אנגלית

^a כזכור, המדד של קלירי לא מתוקן לקיצוץ תחום והמדדים של תנאי הגבול והפרש d-ים מתוקנים לקיצוץ תחום.

הנתונים בלוח 18 מעידים שגם במחקר 2 המתאמים הגבוהים ביותר הם בין המדד של קלירי למדד של הפרש d-ים (בממוצע מעבר לחזאים: 0.61), והמתאמים בין המדד של תנאי הגבול למדד של הפרש d-ים דומים להם (בממוצע: 0.58). אך במחקר זה, להבדיל ממה שנמצא במחקר 1, המתאמים בין המדד של תנאי הגבול למדד של קלירי אינם נמוכים (בממוצע: 0.48). כמו במחקר 1, כאשר כל המדדים אינם מתוקנים לקיצוץ תחום, עולים המתאמים ביניהם, וכאשר המתאמים בין המדדים (כולם ללא תיקון לקיצוץ תחום) מחושבים רק בחוגים שבהם נמצאה הטיה על פי תנאי הגבול (שלוש העמודות השמאליות בלוח 19), מוצאים מתאמים גבוהים עוד יותר³².

בהשוואה בין המחקרים ניתן להתרשם שהמתאמים בין מדדי ההטיה במחקר 2 גבוהים מהמתאמים שהתקבלו במחקר 1.

³¹ כצפוי, השינוי במתאמים ניכר בעיקר ביחס למתאמים של המדד של תנאי הגבול עם שני המדדים האחרים.
³² כצפוי, השינוי במתאמים ניכר בעיקר ביחס למתאמים של המדד של תנאי הגבול עם שני המדדים האחרים.

נספח 4

בדיקת האפקט המשולב של שפת היבחנות ומגדר

בבדיקה זו נכללו חוגים שיש בהם לפחות 2 סטודנטים מכל צירוף של שפת היבחנות ומגדר. לצורך זה הוגדרו ארבע קבוצות: גברים נבחני עברית, גברים נבחני ערבית, נשים נבחנות עברית ונשים נבחנות ערבית. כמו במחקרים 1 ו-2 נבדקו גם כאן ההיבט של תוקף דיפרנציאלי וההיבט של הטיה בבררה. ההיבט של הטיה בבררה נבדק באמצעות מודל הרגרסיה של קלירי. מספר הסטודנטים לפי תחומי לימוד וקבוצה ומספר החוגים לפי תחומי לימוד מוצגים בלוח 20.

לוח 20

מספר הסטודנטים ומספר החוגים בבדיקת האפקט המשולב של שפת היבחנות ומגדר

סה"כ	תחומי לימוד		
	מילוליים	כמותיים	
10155	4324	5831	סטודנטים עברית גברים
2446	853	1593	סטודנטים ערבית גברים
13471	7308	6163	סטודנטים עברית נשים
4088	2528	1560	סטודנטים ערבית נשים
255	102	153	חוגים

בלוח 21 מוצגים סטטיסטיים תיאוריים של החזאים והקריטריון.

לוח 21

ממוצעים וסטיות תקן של החזאים והקריטריון לפי שפת היבחנות, מגדר ותחומי לימוד

תחומי לימוד	משתנה	עברית גברים		עברית נשים		ערבית נשים	
		ממוצע	סטיית תקן	ממוצע	סטיית תקן	ממוצע	סטיית תקן
מילוליים	צש"א	83.5	7.5	71.8	10.7	73.4	9.7
	סכס	55.2	7.6	54.1	6.2	55.5	5.9
	בגרות	93.0	7.5	98.9	6.8	102.4	6.3
	מכפ"ל	607.7	66.9	538.0	51.7	525.9	50.4
	מילולי	118.3	14.1	106.2	12.4	105.7	12.7
	כמותי	118.6	13.4	112.2	12.4	107.0	11.4
	אנגלית	120.7	17.5	96.3	14.4	96.9	13.1
	צש"א	80.2	10.2	70.5	12.0	69.3	11.8
כמותיים	סכס	61.6	5.3	62.5	4.0	62.9	4.1
	בגרות	101.3	6.3	108.8	4.4	110.9	3.6
	מכפ"ל	675.4	41.6	623.3	37.7	610.8	41.5
	מילולי	127.8	10.2	120.1	11.0	120.2	11.6
	כמותי	132.6	9.0	129.4	8.8	124.5	10.0
	אנגלית	133.3	12.3	109.2	14.8	107.8	12.6
	צש"א	81.8	9.0	71.1	11.3	71.3	10.8
	סכס	58.4	6.5	58.4	5.3	59.2	5.1
כולם	בגרות	97.2	6.9	103.8	5.7	106.6	5.1
	מכפ"ל	641.7	55.7	580.8	45.2	568.5	46.1
	מילולי	123.0	12.3	113.2	11.7	113.0	12.2
	כמותי	125.6	11.4	120.8	10.7	115.8	10.7
	אנגלית	127.0	15.1	102.8	14.6	102.4	12.9
	צש"א	81.8	9.0	71.1	11.3	71.3	10.8

מקדמי המתאם בין החזאים לקריטריון, מתוקנים לקיצוץ תחום ונצפים, מוצגים בלוחות 22 ו-23 בהתאמה.

לוח 22

מקדמי התוקף (מתוקנים לקיצוץ תחום) של החזאים לפי שפת היבחנות, מגדר ותחומי לימוד

תחומי לימוד	חזאי	עברית גברים	עברית נשים	ערבית גברים	ערבית נשים
מילוליים	סכם	0.45	0.28	0.53	0.40
	בגרות	0.40	0.22	0.49	0.35
	מכפ"ל	0.36	0.22	0.45	0.36
	מילולי	0.34	0.15	0.39	0.24
	כמותי	0.30	0.21	0.38	0.31
	אנגלית	0.19	0.17	0.32	0.27
כמותיים	סכם	0.52	0.32	0.53	0.37
	בגרות	0.50	0.29	0.51	0.35
	מכפ"ל	0.38	0.25	0.43	0.31
	מילולי	0.27	0.17	0.34	0.19
	כמותי	0.37	0.21	0.41	0.32
	אנגלית	0.23	0.23	0.28	0.20
כולם	סכם	0.49	0.30	0.53	0.39
	בגרות	0.46	0.26	0.50	0.35
	מכפ"ל	0.37	0.24	0.44	0.34
	מילולי	0.30	0.16	0.36	0.22
	כמותי	0.34	0.21	0.40	0.32
	אנגלית	0.21	0.21	0.30	0.24

לוח 23

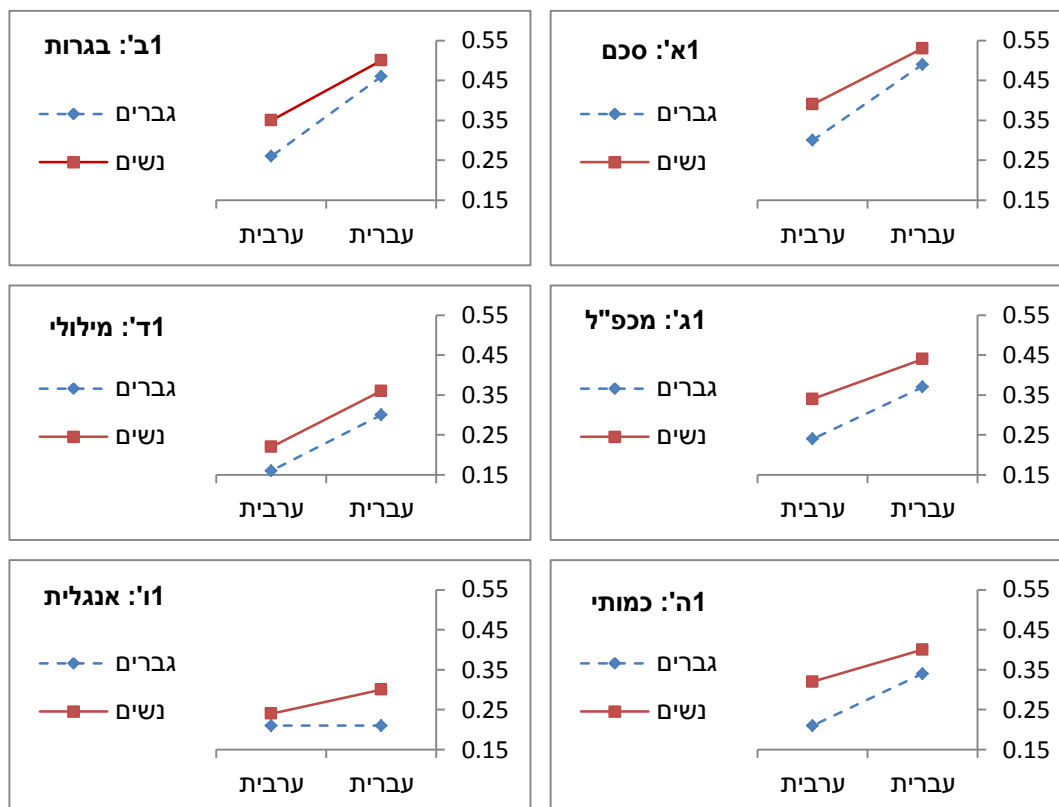
מקדמי תוקף נצפים של החזאים לפי שפת היבחנות, מגדר ותחומי לימוד

תחומי לימוד	חזאי	עברית גברים	עברית נשים	ערבית גברים	ערבית נשים
מילוליים	סכם	0.36	0.21	0.41	0.29
	בגרות	0.31	0.14	0.35	0.24
	מכפ"ל	0.25	0.14	0.29	0.24
	מילולי	0.25	0.09	0.25	0.14
	כמותי	0.21	0.14	0.24	0.20
	אנגלית	0.12	0.12	0.18	0.17
כמותיים	סכם	0.33	0.18	0.34	0.24
	בגרות	0.31	0.16	0.31	0.24
	מכפ"ל	0.14	0.08	0.18	0.12
	מילולי	0.05	0.01	0.09	-0.01
	כמותי	0.21	0.12	0.22	0.19
	אנגלית	0.04	0.08	0.06	0.10
כולם	סכם	0.34	0.19	0.38	0.27
	בגרות	0.31	0.15	0.33	0.24
	מכפ"ל	0.19	0.10	0.24	0.19
	מילולי	0.14	0.04	0.18	0.08
	כמותי	0.21	0.13	0.23	0.19
	אנגלית	0.07	0.09	0.13	0.14

בתרשימים 1א' עד 1ו' מוצגים מקדמי התוקף של החזאים (מתוקנים לקיצוץ תחום) לפי שפת היבחנות ומגדר (מעבר לכל החוגים).

תרשימים 1א' עד 1ו'

מקדמי התוקף של החזאים לפי שפת היבחנות ומגדר (מעבר לכל החוגים)



ניתן להיווכח ששני הקווים בכל תרשים מקבילים זה לזה בקירוב. במילים אחרות, האפקט של מגדר דומה בין שתי הקבוצות של שפת היבחנות (ובהתאם גם האפקט של שפת היבחנות דומה בין שתי קבוצות המגדר). המסקנה היא שלא מתקיימת אינטראקציה בקשר בין שפת היבחנות ומגדר לתוקף הניבוי. באופן מפורט, בתוך כל מגדר תוקף כל החזאים³³ בקרב נבחני עברית גבוה מן התוקף בקרב נבחני ערבית, ובתוך שפת היבחנות תוקף כל החזאים בקרב נשים גבוה מן התוקף בקרב גברים (נציין שבמחקר 2 נמצא שתוקף הבגרות גבוה יותר בקרב הגברים. הבדל זה בין האפקט העיקרי של מגדר לבין האפקטים הפשוטים שלו מוסבר בחלקו על ידי ההבדל בין המדגמים של מחקר 2 והבדיקה הנוכחית וכן על ידי ההבדלים בין המגדרים בפרופורציות של נבחני עברית ונבחני ערבית).

בלוח 24 מוצג המדד של קלירי להטיה בבררה כלפי ארבע הקבוצות.

³³ להוציא את אנגלית שהתוקף שלו בקרב גברים משתי שפות היבחנות הוא דומה.

הטיה בבררה לפי המדד של קלירי^a כלפי קבוצות המוגדרות לפי שפת היבחנות ומגדר

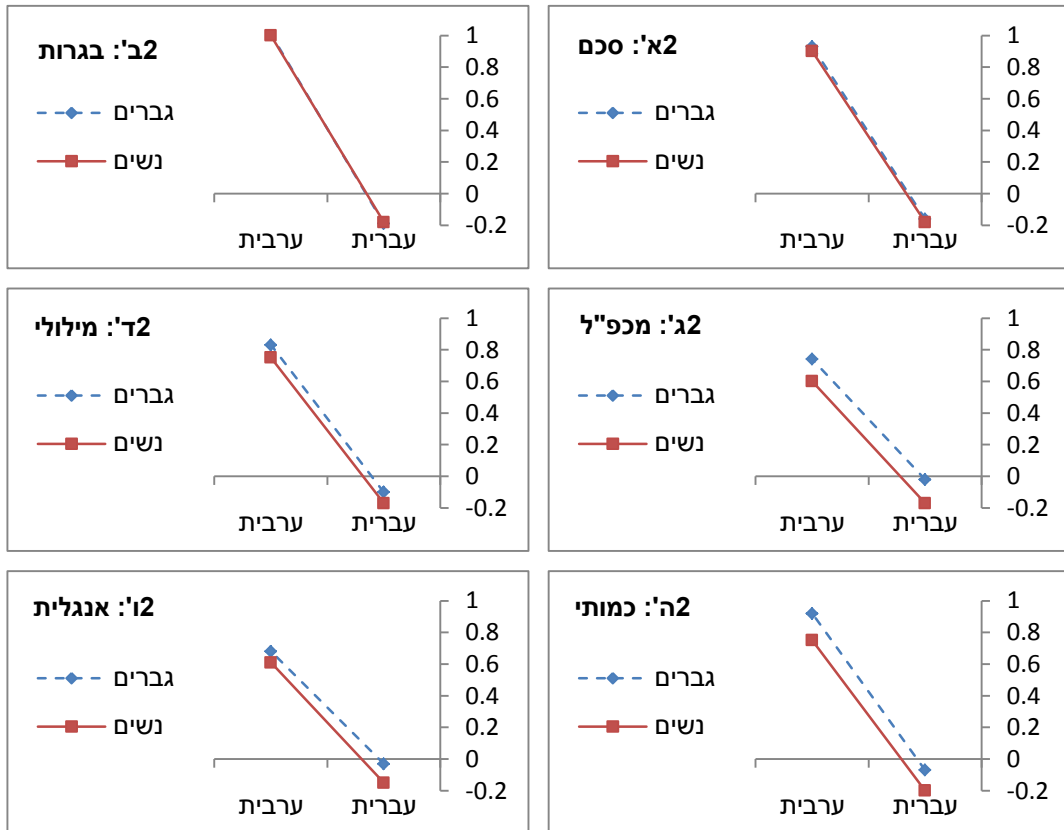
תחומי לימוד	חזאי	עברית גברים	עברית נשים	עברית גברים	עברית נשים
מילוליים	סכום	-0.17	1.12	-0.25	0.96
	בגרות	-0.21	1.17	-0.25	1.04
	מכפ"ל	0.01	0.94	-0.23	0.66
	מילולי	-0.07	0.98	-0.22	0.77
	כמותי	-0.06	1.13	-0.27	0.84
	אנגלית	-0.02	0.85	-0.21	0.66
כמותיים	סכום	-0.16	0.74	-0.11	0.84
	בגרות	-0.17	0.85	-0.11	0.96
	מכפ"ל	-0.05	0.54	-0.12	0.54
	מילולי	-0.12	0.69	-0.12	0.74
	כמותי	-0.08	0.71	-0.15	0.66
	אנגלית	-0.04	0.50	-0.09	0.55
כולם	סכום	-0.16	0.93	-0.18	0.90
	בגרות	-0.19	1.01	-0.18	1.00
	מכפ"ל	-0.02	0.74	-0.17	0.60
	מילולי	-0.10	0.83	-0.17	0.75
	כמותי	-0.07	0.92	-0.21	0.75
	אנגלית	-0.03	0.68	-0.15	0.61

^a במונחי סטיות תקן של הקריטריון; ערך חיובי מבטא הטיה לטובת הקבוצה וערך שלילי הטיה לרעתה.

בתרשימים 2א' עד 2ו' מוצגים ערכי ההטיה בבררה של החזאים כלפי ארבע הקבוצות המוגדרות לפי שפת היבחנות ומגדר (מעבר לכל החוגים).

תרשימים 2א' עד 2ו'

הטיה בבררה לפי המדד של קלירי כלפי קבוצות המוגדרות לפי שפת היבחנות ומגדר



ניתן להיווכח ששני הקווים בכל תרשים מקבילים זה לזה. במילים אחרות, האפקט של מגדר דומה בשתי הקבוצות של שפת היבחנות (ובהתאם גם האפקט של שפת היבחנות דומה בין שתי קבוצות המגדר). המסקנה היא שלא מתקיימת אינטראקציה בקשר בין שפת היבחנות ומגדר לבין הטיה בבררה. באופן מפורט, בתוך כל מגדר יש הטיה לטובת נבחני ערבית. האפקט של מגדר בתוך כל שפת היבחנות אינו אחיד בין החזאים: בבגרות ובציון הסכס אין הטיה לפי מגדר (נציין שבמחקר 2 נמצא שבגרות, ובמידה פחותה גם ציון הסכס, מוטים לטובת נשים. הבדל זה בין האפקט העיקרי של מגדר לבין האפקטים הפשוטים שלו מוסבר בחלקו על ידי ההבדל בין המדגמים של מחקר 2 והבדיקה הנוכחית וכן על ידי ההבדלים בין המגדרים בפרופורציות של נבחני ערבית ונבחני ערבית) ואילו במכפ"ל ובמרכיביו יש הטיה מזערית לרעת נשים. ניתן להיווכח שוב שהאפקט של שפת היבחנות על ממדי ההטיה גדול בהרבה מן האפקט של מגדר.