

הוגנות מערכת המיון להשכלה הגבוהה
כלפי מבקשי תנאים מותאמים
בבחינה הפסיכומטרית (בשפה העברית)

דביר קלפר
אליוט טורוול
תמר קנת-כהן
כרמל אורן

ינואר 2013



דוח מרכז 386
ISBN:978-965-502-175-2

הוגנות מערכת המיון להשכלה הגבוהה
כלפי מבקשי תנאים מותאמים
בבחינה הפסיכומטרית (בשפה העברית)

דביר קלפר, אליוט טורוול, תמר קנת-כהן וכרמל אורן

ינואר 2013

תוכן עניינים

5	1. תמצית
6	2. מבוא
6	2.1. בחינות מותאמות
7	2.1.1. אוכלוסיית מבקשי ההתאמות
8	2.1.2. ההתאמות הניתנות במדור בחינות מותאמות במאל"ו
9	2.2. מטרת המחקר
10	3. שיטה
10	3.1. בניית קובץ הנתונים
10	3.1.1. מקבלי התאמות
11	3.1.2. מסורבי התאמות
11	3.1.3. קובץ התוקף
12	3.1.4. אוכלוסיית המחקר
13	3.2. משתני המחקר
14	3.3. עיבוד הנתונים
14	3.3.1. מחקר מסורתי
15	3.3.2. מחקר בשיטת ציוני נמייה
19	4. תוצאות
21	4.1. תוצאות מפורטות
21	4.1.1. מחקר מסורתי
22	4.1.2. מחקר בשיטת ציוני נמייה
23	4.2. סיכום תוצאות
25	5. דיון
25	5.1. הטיה בבדיקה
27	5.2. תוקף דיפרנציאלי
28	5.3. השוואה בין השיטות
28	5.3.1. ה- Δ של כהן
30	5.3.2. מדד קלירי להונות
32	5.3.3. תוקף דיפרנציאלי
35	6. מקורות
39	7. נספחים
39	7.1. מודלים ושיטות
39	7.1.1. ה- Δ של כהן
39	7.1.2. המודל של קלירי
40	7.1.3. מקדם תוקף
40	7.1.4. ציוני נמייה
41	7.2. הרחבה על שתי השורות הראשונות בלוח 18

רשימת לוחות

- 10..... לוח 1 - קידוד מקורי ללקויות וחלוקתן לשלוש קטגוריות
- 12..... לוח 2 - שכיחות לומדים לפי שנת תחילת הלימודים
- 13..... לוח 3 - התפלגות האוכלוסייה
- 15..... לוח 4 - שכיחות הלומדים בכל קבוצה בהשוואות השונות
- 16..... לוח 5 - ממוצעים וסמיות תקן עבור משתני הרקע מצב כלכלי, השכלת אב והשכלת אם
- 16..... לוח 6 - ממוצעים וסמיות תקן עבור משתני הרקע מין וגיל
- לוח 7 - ממוצעים וסמיות תקן עבור משתני הרקע מצב כלכלי, השכלת אב והשכלת אם
- 17..... לאחר השלמת נתונים חסרים
- 18..... לוח 8 - גודל ההשוואה עם יחידות העיבוד הרלוונטיות לפני ואחרי הזיווג
- 19..... לוח 9 - ממוצע וסמיית תקן עבור הקריטריון והחזאים
- 19..... לוח 10 - ה- Δ של כהן עבור חזאים מרכזיים – מחקר מסורתי
- 20..... לוח 11 - ה- Δ של כהן עבור חזאים משניים – מחקר מסורתי
- 20..... לוח 12 - ה- Δ של כהן עבור חזאים מרכזיים – מחקר ציוני נטייה
- 20..... לוח 13 - ה- Δ של כהן עבור חזאים משניים – מחקר ציוני נטייה
- 21..... לוח 14 - מדדי המיה* ביחס לקבוצת ההתייחסות על פי קלירי – מחקר מסורתי
- 21..... לוח 15 - מקדם תוקף* – מחקר מסורתי
- 22..... לוח 16 - מדדי המיה* ביחס לקבוצת ההתייחסות על פי קלירי – מחקר ציוני נטייה
- 22..... לוח 17 - מקדם תוקף* – מחקר ציוני נטייה
- 23..... לוח 18 - סיכום תוצאות: חזאים מרכזיים
- 23..... לוח 19 - סיכום תוצאות: חזאים משניים
- 23..... לוח 20 - מקרא
- 41..... לוח 21 - ה- Δ של כהן עבור חזאים מרכזיים ושתי הקבוצות הראשונות
- 42..... לוח 22 - התפלגות מין וגיל לפי קבוצות (0, 1, 2-1)
- 42..... לוח 23 - ממוצע וסמיית תקן של הקריטריון ושני החזאים המרכזיים לפי גיל
- 42..... לוח 24 - ממוצע וסמיית תקן של הקריטריון ושני החזאים המרכזיים לפי מין

רשימת איורים

- 28..... איור 1 - ה- Δ של כהן עבור קבוצה 1 = לקויות למידה
- 28..... איור 2 - ה- Δ של כהן עבור קבוצה 2 = קשב וריכוז
- 29..... איור 3 - ה- Δ של כהן עבור קבוצה 3 = מגבלות פיזיות
- 29..... איור 4 - ה- Δ של כהן עבור קבוצה 98 = "חסרי נתונים"
- 29..... איור 5 - ה- Δ של כהן עבור קבוצה 99 = מסורבים מקצועית
- 30..... איור 6 - מדד קלירי עבור קבוצה 1 = לקויות למידה
- 30..... איור 7 - מדד קלירי עבור קבוצה 2 = קשב וריכוז
- 31..... איור 8 - מדד קלירי עבור קבוצה 3 = מגבלות פיזיות
- 31..... איור 9 - מדד קלירי עבור קבוצה 98 = "חסרי נתונים"
- 31..... איור 10 - מדד קלירי עבור קבוצה 99 = מסורבים מקצועית
- 32..... איור 11 - מקדם תוקף עבור קבוצה 1 = לקויות למידה
- 32..... איור 12 - מקדם תוקף עבור קבוצה 2 = קשב וריכוז
- 33..... איור 13 - מקדם תוקף עבור קבוצה 3 = מגבלות פיזיות
- 33..... איור 14 - מקדם תוקף עבור קבוצה 98 = "חסרי נתונים"
- 34..... איור 15 - מקדם תוקף עבור קבוצה 99 = מסורבים מקצועית

1. תמצית

במחקר זה נבדקה הוגנות מערכת המיון לאוניברסיטאות כלפי מבקשי תנאים מותאמים בבחינה הפסיכומטרית. בפרט, מחקר זה עוסק הן באלו שביקשו וקיבלו התאמות, הן בנבחנים שבקשתם סורבה אם מסיבות טכניות אם מסיבות מקצועיות.

מבקשי ההתאמות חולקו לפיכך לחמש קבוצות: שלוש קבוצות שקיבלו התאמות – בעלי לקויות למידה, בעלי בעיות קשב וריכוז ובעלי מגבלות פיזיות – ושתי קבוצות שבקשתם לקבל התאמות נדחתה – האחת בשל מסיבות טכניות (חוסר נתונים), והשנייה מנימוקים מקצועיים.

שאלת ההוגנות נבדקה משני היבטים: ההיבט של הטיה בברירה וההיבט של תוקף דיפרנציאלי. לצורך בדיקת הטיה בברירה נעשה שימוש במודל הרגרסיה של קלירי (Cleary, 1968), שבדק ניבוי דיפרנציאלי. השימוש במודל זה, המביא בחשבון את מגבלות התוקף של החזאי, מומלץ מבחינה מקצועית (AERA/APA/NCME, 1999) והוא השיטה המקובלת ביותר לגילוי ואיתור הטיה בברירה (Young, 2001). לצורך בדיקת התוקף הדיפרנציאלי נעשה שימוש במקדמי המתאם.

נערכו שני מחקרים השונים ביניהם באופן הגדרת קבוצת ההתייחסות (נבחנים בתנאים רגילים), שאליה מושוות קבוצות המוקד. המחקר הראשון בוצע באופן מסורתי כפי שנעשה עד היום במחקר מעין זה – קבוצת ההתייחסות היא כלל הנבחנים ה"רגילים". מחקר זה מאפשר להשוות את תוצאות המחקר הנוכחי לתוצאות שהתקבלו בעבר. המחקר השני משתמש בשיטת ציוני הנטייה על מנת לבחור, בעבור כל קבוצת מוקד (שחבריה ביקשו התאמות), תת-קבוצה מתוך קבוצת ה"רגילים" שהיא בעלת אפיוני רקע דומים לאלו של קבוצת המוקד.

המחקר בוצע על נתוני 124,501 סטודנטים מכל האוניברסיטאות בישראל, אשר למדו בשנה א' במחזורים 2002/3 עד 2008/9 ב-2,036 חוגים שונים ואשר נבחנו בבחינה הפסיכומטרית החל ממועד יולי 2000 בשפה העברית.

בשני המחקרים נמצאו תוצאות זהות כמעט לגבי שאלת התוקף הדיפרנציאלי. בשאלת ההטיה בברירה התקבלו בשני המחקרים תוצאות דומות, כאשר גודלי האפקט היו קטנים יותר בשיטת ציוני הנטייה.

לא נמצאה הטיה בברירה לטובת או לרעת בעלי לקויות למידה, בעלי בעיות קשב וריכוז ואלה שבקשתם נדחתה מנימוקים מקצועיים; נמצאה הטיה קלה לטובת קבוצת בעלי המגבלות הפיזיות (ניבוי יתר) ולרעת קבוצת הנבחנים שבקשתם נדחתה בשל חוסר נתונים (ניבוי חסר).

מקדמי התוקף של ציוני הבחינה הפסיכומטרית היו זהים עבור קבוצת ההתייחסות (קבוצת ה"רגילים" שלא ביקשו התאמות בהתאם להשוואה שבוצעה) וקבוצות המוקד (חמש קבוצות המחקר). לעומת זאת, מקדמי התוקף של הבגרות וציון הסכם היו גבוהים יותר עבור קבוצת ההתייחסות מאשר עבור קבוצות המוקד (פרט ל"חסרי הנתונים"), במיוחד, בקרב בעלי לקויות למידה ובעלי בעיות קשב וריכוז.

2.1. בחינות מותאמות

התאמות מיוחדות לנבחנים בעלי לקויות נעשו עם השנים דבר שבשגרה עבור מבחנים הנערכים בקנה מידה גדול ואשר מהווים מבחנים חורצי גורלות. תהליך זה מסייע בהעלאת הפרופורציה של סטודנטים בעלי לקויות בהשכלה הגבוהה. יתכן כי הלכות/המגבלה יכולות למנוע מהמועמד הזדמנות שווה להציג את הידע או היכולת אשר נמדדים בבחינה (Willingham et al., 1998). הסוגיה של התאמות מיוחדות מעלה מגוון של שאלות, כמו: מהי הדרך הנאותה לקבוע אלו לקויות מצדיקות בחינה בתנאים מותאמים? מהי התאמה הולמת והוגנת כלפי לקות מסוימת? באילו אמות מידה יש להשתמש כדי לקבוע זאת? איך לפרש את הציון כך שימנע ניבוי יתר/חסר של ההצלחה האקדמית, ובכך למנוע הליך קבלה לא הוגן (Pitoniak & Royer, 2001)?

התאמות הן שינוי באמות המידה התקניות שמטרתן להפוך את "מגרש המשחקים" למאוזן יותר כלפי האוכלוסייה בעלת הלקויות על ידי הפחתת השונות הלא רלוונטית הנגרמת מהלקות. ציונים שהושגו על בסיס בחינה מותאמת שניתנה בצורה תקפה מטרתם למדוד בקרב נבחנים בעלי לקויות אותה התכונה שנמדדת על ידי מבחנים תקינים בקרב האוכלוסייה ה"רגילה" (Tindal & Fuchs, 1999; Lovett, 2010).¹

למרות העניין הרב בנוגע לתוקף המבחן ופירוש התוצאות, בארץ לא נעשה מחקר על מנת לבחון את השפעות ההתאמות על תוצאות המבחן, למעט במרכז הארצי לבחינות ולהערכה² (אנגל, 2007; Oren & Even, 2005). הקושי המרכזי בביצוע מחקרים כאלו נובע בעיקר מן העובדה שקשה למצוא קבוצות של אנשים עם לקויות שיכולות להוות יחידות עיבוד מספיק גדולות על מנת לבצע סוג כזה של מחקר.

נכון להיום, מחקרים שבדקו את הוגנות ההתאמות התייחסו לשני היבטים. ההיבט הראשון הוא תוקף ניבוי (כלומר, מהי מידת הדמיון של מקדמי תוקף הניבוי בין קבוצת ההתייחסות ובין קבוצת המוקד). השני הוא יכולת השוואת הציונים (כלומר, עד כמה דומים ציוני הקריטריון בקבוצת ההתייחסות ובקבוצת המוקד, כשציוני החזאי מוחזקים קבוע).

מחקרים בנושא תוקף ניבוי מראים שבדרך כלל ציונים המושגים בעזרת התאמות הם בעלי מתאם נמוך יותר (בהשוואה למתאם של ציונים שמושגים ללא התאמות) עם הקריטריון (Laing & Farmer, 1984; Braun, Ragosta & Kaplan, 1986; Ziomek & Andres, 1996). הקריטריון שבו משתמשים בדרך כלל הוא ציונו הממוצע של הסטודנט באוניברסיטה בשנת הלימודים הראשונה.

¹ בהקשר זה ראה גם את חוק זכויות תלמידים עם לקות למידה במוסדות על-תיכוניים, התשס"ח-2008.

² להלן מאל"ן. www.knesset.gov.il/privatelaw/data/17/3/153_3_1.rtf

מחקרים בנושא של יכולת השוואת הציונים מראים שציונים של נבחנים בעלי לקויות שקיבלו התאמות בבחינה דומים לאלו של נבחנים שנבחנו ללא התאמות. ישנן עדויות מחקריות לכך שהתאמות יכולות לעזור ליחידים עם לקויות בהשגת ציונים בני-השוואה לציונים של יחידים ללא לקויות (Centra, 1986; Bennet, Rock, & Jurele, 1986; Willingham et al., 1988; Ziomek & Andrews, 1996). כמה מחקרים התמקדו בשאלת הקשר בין הארכת זמן (התאמה מבוקשת) לבין תוצאות הבחינה. במחקרים אלה נמצא מתאם נמוך בין תוספת זמן וציון ממוצע לתואר (Braun, Ragosta, & Kaplan, 1986; Ziomek & Andrews, 1996; Zurcher & Bryant, 2001).

2.1.1. אוכלוסיית מבקשי ההתאמות

נבחנים בעלי לקויות אינם אוכלוסייה הומוגנית, אלא מאופיינים על ידי ספקטרום רחב של בעיות מוטוריות, חושיות וקוגניטיביות. בעבר, התאמות ניתנו בעיקר לנבחנים בעלי נכות פיזית או חושית. בשנים האחרונות, עם גידול במודעות וההבנה כלפי קבוצת האנשים בעלי קשיי למידה, דגש רב יותר מושם על הצורך במתן התאמות לנבחנים בעלי לקויות קוגניטיביות. אכן, קבוצת לקויי הלמידה היא כיום הקבוצה הגדולה ביותר הפונה בבקשה למתן התאמות במבחנים (Camara, 1998; Wightman, 1993). עובדה זו שבה ומעלה את אותן השאלות בנוגע לשינוי התקפות של תוצאות המבחן (Philips, 1994).

2.1.1.1. לקות למידה

לקות למידה מוגדרת כ"הפרעה באחד או יותר מהתהליכים הפסיכולוגיים הבסיסיים המעורבים בהבנה או בשימוש בלשון שבאה לידי ביטוי בעיקר ביכולת לא מושלמת בהאזנה, חשיבה, דיבור, קריאה, כתיבה, איות או בביצוע של חישובים מתמטיים" (U.S Department of Education, Guideline, 1996, cited in Pitoniak & Royer, 2001, p.58). בעבר, האבחון של לקות למידה נעשה על ידי השוואת רמת ההישגים המתקבלת במבחן מסוים עם הרמה הצפויה בהתבסס על מבחני אינטליגנציה (Gresham, MacMillan, & Bocian, 1996). בעקבות בעיות מושגיות ומעשיות הקשורות למודל זה, מקובלים כיום מודלים אחרים להגדרת לקות למידה. למשל, מודלים חדשים (Flanagan, Fiorello, & Ortiz, 2010) קובעים שיש להעריך את ההתאמה בין תפקודים אקדמיים ותפקודים קוגניטיביים הקשורים בקשר רלוונטי לתפקודים האקדמיים ויכולות אינטלקטואליות מגוונות. עקב כך גם ההגדרות ב-DSM-IV-TR, (APA, 2000) עתידות להשתנות (Scanlon, 2013). לקות למידה תוגדר כמצב של הלימה בין תפקודים אקדמיים נמוכים ומנגנונים קוגניטיביים רלוונטיים פגועים בתוך הקשר של יכולת אינטלקטואלית תקינה או למעלה מכך. לעיתים קרובות נעשה שימוש גם בקריטריון של תגובה להוראה (Fletcher et al., 2002), כאשר לקויי למידה מוגדרים כמי שלא הגיבו להוראה איכותית אליה מגיבים רב התלמידים.

ההתאמות המאושרות בדרך כלל לנבחנים בעלי לקויות למידה כוללות הארכת משך הבחינה, שימוש בהקראה, שכתוב והקלדת התשובות ושימוש בעזרי קריאה וכתובה. השימוש בהתאמות אלו נועד לעקוף פגמים ביכולת עיבוד המידע ומגבלות פיזיות שונות אצל נבחנים עם לקות למידה (MacArthur, 1996).

2.1.1.2 הפרעת קשב, ריכוז והיפראקטיביות

הפרעת קשב, ריכוז והיפראקטיביות (באנגלית: ADHD; נקראת בקיצור הפרעת קשב וריכוז) היא הפרעה הנובעת כנראה מהתפתחות פתולוגית של מערכת העצבים המרכזית. תסמיני הליבה של הפרעת קשב וריכוז הם: קשיי קשב וריכוז, היפראקטיביות ואימפולסיביות. עם זאת, לא כל התסמינים חייבים להופיע כדי לענות על הקריטריונים של אבחנת הפרעת קשב וריכוז. (ADD) הוא מונח ששימש בעבר לתיאור הפרעת קשב וריכוז ללא היפראקטיביות.

הפרעת קשב וריכוז יכולה לבוא לידי ביטוי בכל תחומי החיים. היא עלולה להפריע בלמידה, בעבודה וביחסים בין-אישיים. אף על פי שהתפקוד האקדמי עלול להיפגע בעקבות הפרעת קשב וריכוז, היא אינה מוגדרת על ידי המדריך האבחוני DSM כלקות למידה; זאת משום שהתפקוד בתחום זה לא בהכרח מושפע מהפרעת הקשב והריכוז ומשום שהיא יכולה לפגוע גם בתחומים שאינם אקדמיים. עם זאת, יש המכלילים גם מיומנויות הקשבה כתחומי למידה שנפגעים בלקות למידה ולפיכך הם כוללים בהגדרת לקויות הלמידה גם קשיים כגון הפרעת קשב וריכוז.

2.1.1.3 נכות פיזית

זיהוי אנשים בעלי נכויות פיזיות הוא ללא כל ספק פחות בעייתי ומורכב מזיהוי אנשים בעלי לקויות למידה. המושג נכות פיזית מוגדר בתור מגבלה בתפקוד גופני הנובעת ממום מולד או מפגיעה פיזית או נפשית, עקב מחלה, תאונה, זדון וכדומה. נשים לב לכך שיתכן והנכות הפיזית תשפיע גם על התפקוד הקוגניטיבי של הנכים. בכל מקרה, קיים צורך במתן התאמות לאנשים אלה. במתן התאמות לאוכלוסייה זו, יש להדגיש כי חשוב לקחת בחשבון את הספקטרום הרחב של הנכויות, גם באופי וגם בחומרה (Willingham et al., 1998).

2.1.2 ההתאמות הניתנות במדור בחינות מותאמות³ במאל"ו

הליך הקבלה להשכלה הגבוהה בישראל מבוסס בעיקר על שני אינדיקטורים: מבחן הבגרות ומבחן כניסה פסיכומטרי לאוניברסיטאות⁴. יחידת במו"ת במאל"ו הוקמה כדי לתת מענה לקשייהם של נבחנים עם מגבלות פיזיות רפואיות, מגבלות פסיכיאטריות, לקויות למידה או הפרעות קשב וריכוז. צוות היחידה מורכב מאנשי מקצוע מהתחומים פסיכולוגיה, חינוך, נירופסיכולוגיה ופסיכומטריקה. בנוסף נוהגים להתייעץ לפי הצורך עם מומחים גם מרפואה. תפקיד הצוות הוא להתאים במידת האפשר את תנאי הבחינה למגבלה או ללקות במטרה לעקוף את הקשיים הנובעים ממנה. ההתאמות נקבעות מתוך ניסיון לשמור במידת האפשר על המדידה התקינה ועל ההוגנות כלפי כלל ציבור הנבחנים.

³ להלן במו"ת.

⁴ להלן מכפ"ל.

המדור מטפל בכמה אלפי נבחנים בשנה כאשר בשנים האחרונות ישנו גידול של כ-10 אחוז במספר הפונים מידי שנה. מספר הפונים ב-2011 עמד על 3,636 כאשר מתוכם 3,152 בקשות היו מלאות (ללא מסמכים חסרים). 2,245 בקשות נענו בחיוב (כ-71 אחוז). (על פי נתונים בשנים 2010-2011).⁵

סוגי ההתאמות שניתנות בהתאם לסוג וחומרת הלקות הם :

- התאמה של המדיה שבה מוצגות ההוראות והשאלות. למשל, נבחנים עם בעיות ראייה יכולים לקבל בחינה ממחושבת המושמעת על גבי המחשב ומותאמת למגבלותיהם, הגדלת חוברת הבחינה, שימוש באמצעי הגדלה (טלוויזיה במעגל סגור) והעזרות בהקראת בוחן אנושי.
- התאמה של המדיה בה עונים על השאלות. למשל, נבחנים שמוגבלים מבחינה פיזית יכולים להיעזר בבוחן שמסמן עבורם את התשובות או בגיליון תשובות מוגדל.
- במקרים רבים ניתנת גם תוספת זמן לביצוע המבחן. ראוי לציין שאין מידע אובייקטיבי בכמה זמן יש להאריך את משך הבחינה עבור בעלי הלקויות השונות. ברוב המקרים תוספת הזמן ניתנת בפורמט של בחינה אדפטיבית ממוחשבת (מפע"ם, ראה להלן).
- תנאים מותאמים אשר נבחנים בעלי בעיות קשב וריכוז עשויים לקבל הינם : הגבלת מספר הנבחנים בחדר הבחינה לשמונה בלבד, מתן הפסקות של חמש דקות בין פרקים, שימוש במחשבון כיס ושימוש בתשובון מוגדל.
- אנשים עם בעיות ראייה/שמיעה נזקקים לפעמים להתאמת תכני המבחן. למשל, החלפת פריטים שתלויים ביכולות ויזואליות בפריטים אחרים שאינם דורשים יכולת ויזואלית. נדגיש, החלפת פריט בפריט אחר יכולה לטמון בתוכה שינויים בעלי אפיון אחר מזה שבמקור התכוונו אליו – לכן, יש לנקוט משנה זהירות בשינויים מסוג זה.
- שינויים נוספים כוללים שינויים פיזיים בתנאי הבחינה (למשל שולחנות וכסאות המיועדים לאנשים עם מגבלות פיזיות).

2.2. מטרות המחקר

מטרת המחקר היא לבדוק את הוגנות מערכת המיון לאוניברסיטאות כלפי מועמדים מבקשי תנאים מותאמים. ההוגנות נבדקת משני היבטים :

- הטיה בברירה
- תוקף דיפרנציאלי

לצורך כך נעשה שימוש בשתי שיטות :

- השיטה המסורתית שבה מושווית כל אחת מקבוצות המוקד לקבוצת הרוב כולה
- שיטת ציוני הנטייה שבה דוגמים לכל קבוצת מוקד קבוצת השוואה מתוך הנבחנים ה"רגילים", אשר דומה לקבוצת המוקד בנתוני הרקע שלה

מטרת משנה של המחקר הנוכחי היא להשוות בין שתי השיטות.⁶

⁵ ראה גם אתר האינטרנט של מאל"ו : <https://www.nite.org.il/index.php/he/accommodations.html>
⁶ באמצעות השיטה המסורתית ניתן להשוות את התוצאות ממחקר זה עם תוצאות המחקר שבוצע בעבר.

3.1. בניית קובץ הנתונים

החל מיולי 2000 נכנסה לשימוש בחינת המפע"ם⁷ (לנבחני עברית בלבד). מאחר שרבים מהסטודנטים המבקשים מבחן מותאם נבחנים במפע"ם הוחלט לאסוף נתוני נבחנים החל ממועד זה (מושינסקי, 2000; Cohen et al., 2008; Moshinsky & Kazin, 2004).

בהתאם לכך התקבלו שני קבצים:

- קובץ מקבלי התאמות (סה"כ 21,497 רשומות).
- קובץ מבקשי התאמות שבקשתם לא התקבלה (סה"כ 11,746 רשומות).

3.1.1. מקבלי התאמות

בקובץ זה לכל נבחן יש קוד לקות המקודד לפי 26 קטגוריות (חלקם מקודדים ביותר מקטגוריה אחת). הוחלט לקודד כל נבחן בקטגוריה אחת בלבד: לקויות למידה, בעיות קשב וריכוז או מגבלות פיזיות בהתאם ללוח הבא.

לוח 1 - קידוד מקורי ללקויות וחלוקתן לשלוש קטגוריות

מגבלות פיזיות (3)		בעיות קשב וריכוז (2)		לקויות למידה (1)	
תיאור	לקות	תיאור	לקות	תיאור	לקות
ליקוי בראייה	2			ליקוי למידה	1
עיוורון	3			לקות קריאה עברית	13
נכות פיזית	4			לקות קריאה אנגלית	14
פגיעת ראש	5			לקות בתחום כמותי	15
ליקוי שמיעה	6			עיבוד איטי	16
אפילפסיה	7			לקות כוללת רב תחומית (G.V.)	17
בעיה נפשית	8			לקות שפתית	18
בעיות גב	9	קשב וריכוז	12	לקות למידה לא מילולית (NLD)	19
סכרת	10				
טורט	11				
כתיבה/גרפומוטורי	20				
קרוהן קוליטיס	21				
CP	22				
הזעת יתר	23				
מחלה כרונית	24				
מצב רפואי זמני	25				
אספרגר	26				

⁷ מבחן פסיכומטרי ממוחשב המיועד לאוכלוסיית הנבחנים בתנאים מותאמים. המבחן עונה על צורכיהם של רוב הנבחנים המיוחדים. גרסה מתקדמת יותר של המבחן, המאפשרת השמעה של פרק האנגלית, מועברת לנבחנים בעלי לקויות למידה.

מאחר שיש נבחנים בעלי מספר לקויות, חלקם מקבוצות סיווג שונות, היה צורך בכלל החלטה כיצד לסווג אותם לסוג לקות אחת בלבד. ההחלטה לגבי קבוצת ההשתייכות אליה שייך כל נבחן התקבלה בהתאם לכלל ההחלטה הבא:

- ☐ כדי להיות מסווג בקטגוריה לקויות למידה (1), על הנבחן להיות בעל לפחות אחת מהלקויות ברשימת לקויות למידה (ניתכן בהחלט שיהיה בעל עוד לקויות משתי הרשימות האחרות).
- ☐ כדי להיות מסווג בקטגוריה בעיות קשב וריכוז (2), הפונה חייב להיות בעל לקות זאת ורק זאת.
- ☐ כדי להיות מסווג בקבוצת המגבלות הפיזיות, על הפונה להיות בעל לפחות אחת מהלקויות ברשימת המגבלות הפיזיות וללא שום לקות מרשימת לקויות הלמידה (ניתכן מצב שיהיה גם עם בעיית קשב וריכוז).

3.1.2. מסורבי התאמות

נבחנים שלא קיבלו תנאים חולקו לשתי קבוצות:⁸

- ☐ "חסרי נתונים" – פונים שפנייתם להתאמות נדחתה מסיבות טכניות (נבחנים שפנו מאוחר מידיי בבקשה להיבחן בתנאים מותאמים ואנשים שבפנייתם חסרו מסמכים רלוונטיים משמעותיים ונשלחה בקשה להשלמת הנתונים)⁹ אשר קודדו כ-98.
- ☐ נבחנים שבקשתם נבדקה על ידי הצוות המקצועי ולאחר קריאת כל החומר הוחלט לסרב לבקשתם. אנשים אלו קיבלו את השם "מסורבים מקצועית". קבוצה זו קודדה כ-99¹⁰.

3.1.3. קובץ התוקף

קובץ מקבלי התנאים וקובץ המסורבים שורשרו והתקבל קובץ אחד גדול הנקרא על ידנו **קובץ המיוחדים**. בקובץ כל אדם (המזוהה לפי קוד הבחינה שלו) מופיע רק פעם אחת¹¹. קבוצות 1, 2, 3, ו-99 הן, אם כן, קבוצות זרות.

קובץ המיוחדים שהתקבל הינו קובץ שכולל את פרטי ה**נבחנים**. מאחר שבמחקר תוקף עניינו הוא בסטודנטים, נרצה לבנות קובץ עבורם. לצורך כך קובץ המיוחדים הוצלב עם **קובץ התוקף האוניברסיטאי** (לאחר שנשמרו בו נבחני מכפ"ל בעברית בלבד החל ממועד יולי 2000). ההצלבה בוצעה לפי קוד הבחינה. משמעות ההצלבה לפי משתנה זה היא שלכל תלמיד הוצמד הציון הגבוה ביותר בבחינה הפסיכומטרית מבין כלל ציוני הבחינות שבהן נבחן לפני תחילת הלימודים. סטודנט שלומד ביותר מאשר חוג אחד נספר בהתאם למספר החוגים בהם הוא לומד.

⁸ למעשה ישנה קבוצה נוספת של אנשים (900) שביטלו את בקשתם להיבחן בתנאים מיוחדים. קבוצה זו, לדעתנו, איננה מסורבת ולכן היא הוצאה מרשימת המסורבים.

⁹ במקור מקודדים בקוד 310 ו-320 בהתאמה.

¹⁰ במקור מקודדים בקוד 300.

¹¹ קובץ מקבלי התנאים נבנה כך. קובץ המסורבים לא הכיל כפולים והינו זר לקובץ מקבלי התנאים.

בקובץ החדש שהתקבל – **קובץ התוקף** – הוסף לכל סטודנט משתנה קבוצה בהתאם לכלל הבא :
 מי שנמצא בשני הקבצים (תוקף אוניברסיטאי ומיוחדים) קיבל את הסיווג הקודם (1, 2, 3, 98 ו-99).

מי שנמצא רק בקובץ התוקף האוניברסיטאי קיבל את הסיווג "רגיל" (0).
 מי שנמצא רק בקובץ המיוחדים (ולא נמצא בקובץ התוקף האוניברסיטאי), כנראה שלא המשיך ללמוד באוניברסיטה (או שלא סיים שנה א').

העיבודים בהמשך מתבססים על קובץ התוקף.

3.1.4. אוכלוסיית המחקר

להלן התפלגות האוכלוסייה (מעבר לקבוצות השונות) לפי השנה האקדמית שבה התחילו ללמוד באוניברסיטה.

לוח 2 - שכיחות לומדים לפי שנת תחילת הלימודים

שנה אקדמית של תחילת הלימודים	שכיחות	אחוזים	שכיחות	אחוזים
	מצטברת	מצטברת	מצטברת	מצטברת
2000/1	1,016	0.72	1,016	0.72
2001/2	9,222	6.54	10,238	7.26
2002/3	15,300	10.85	25,538	18.11
2003/4	15,820	11.22	41,358	29.33
2004/5	16,970	12.04	58,328	41.37
2005/6	17,815	12.63	76,143	54.00
2006/7	19,770	14.02	95,913	68.02
2007/8	18,879	13.39	114,792	81.41
2008/9	19,947	14.15	134,739	95.56
2009/10	6,259	4.44	140,998	100.00

כפי שניתן לראות, שנים 2000/1, 2001/2 ו-2009/10 אינן מלאות מבחינת הלומדים. ההסבר לגבי השנתיים הראשונות הוא שכנראה חלק גדול מהאנשים שלמדו בשנים אלו נבחנו לפני יולי 2000. ההסבר לגבי שנת 2009/10 הוא שקובץ התוקף האוניברסיטאי שעמד לרשותנו לא היה מלא לשנה זו.

בהתאם לכך, אוכלוסיית המחקר כוללת את הנבחרים בעברית שהתחילו את לימודיהם באוניברסיטאות בשנים 2002/3 עד 2008/9. בנוסף, נלקחו רק אנשים שעבורם נמצא ציון שנה א' וציון בגרות מעל 10 (מתוך 100)¹² (וזאת בנוסף לתנאים הקודמים של נבחרים החל מיולי 2000 בשפה העברית).

התפלגות האוכלוסייה שהתקבלה לפי קבוצות המחקר היא כדלהלן :

¹² זאת כדי להימנע מנתונים שהוקלדו באופן שגוי; למשל, ציון בגרות של 70 שהוקלד כ-7.

	קבוצה	שכיחות	אחוזים
לא ביקשו התאמות	0 = רגילים	120,503	96.79
	1 = לקויות למידה	958	0.77
ביקשו התאמות וקיבלו	2 = קשב וריכוז	187	0.15
	3 = מגבלות פיזיות	1,096	0.88
ביקשו התאמות ולא קיבלו	98 = "חסרי נתונים"	299	0.24
	99 = מסורבים מקצועית	1,458	1.17

על אוכלוסייה זו ביצענו תיקנון של החזאים והקריטריון (בהתאם לסוג המחקר) ברמת יחידת עיבוד המוגדרת כחוג במוסד בשנת לימודים.

3.2. משתני המחקר

חזאים

1. בגרות – ממוצע משוקלל במספר היחידות של ציוני תעודת הבגרות; ציונים אלו הינם ממוצע בין ציוני המגן וציוני בחינת הבגרות. סולם: 10-120 (החישוב כולל בונוסים שונים עבור מקצועות מתוגברים).
2. הציון הכללי במכפ"ל – המבחן מורכב משלושה חלקים (ראה חזאים 4 עד 6 להלן) ומשוקלל לפי 40% לחלק המילולי, 40% לחלק הכמותי ו-20% לאנגלית. מכפ"ל הוא סוללת מבחנים רבי-ברירה, המיועד למדוד אספקטים שונים של יכולות לימודיות ומטרתו לחזות הישגים אקדמיים בעתיד. סולם הציונים של מכפ"ל נע בין 200 ו-800, עם ממוצע בסיסי 500 וסטיית תקן 100.
3. ציון סכס – חושב כצירוף במשקלות שווים של מכפ"ל ובגרות בקרב המועמדים לכל מוסד. סולם ציון הסכס נע בין 0 ו-100, עם ממוצע 50 וסטיית תקן 10¹³.
4. הציון בתחום חשיבה מילולית במכפ"ל – חלק זה מכיל 60 פריטים ומתמקד בכושר וביכולות המילוליות המופיעים בלימודים האקדמיים: אנליזה והבנה של חומר כתוב מורכב, חשיבה שיטתית ולוגית, הבחנה בין משמעויות של מילים ומושגים. החלק המילולי מכיל פריטים כמו מילים נרדפות, דברים והיפוכם, הקשים, השלמת משפטים, לוגיקה והבנת הנקרא.
5. הציון בתחום חשיבה כמותית במכפ"ל – חלק זה מכיל 50 פריטים שמתמקדים בשימוש במספרים ובמושגים מתמטיים (אלגבריים וגיאומטריים) בפתרון בעיות כמותיות וניתוח

¹³ באופן מפורט, ציון הסכס מחושב 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 הסטטיסטיקה הני"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. פרוצדורה זאת מהווה קירוב לציון הקבלה אותו מחשבות האוניברסיטאות.

מידע המוצג בגרפים, טבלאות ותרשימים. בחלק זה נדרש רק ידע בסיסי במתמטיקה – כזה שנרכש בכיתה ט' או י' במרבית מוסדות החינוך בישראל. נוסחאות והסברים של מונחים מתמטיים שאליהם יתכן ויזדקק הנבחן במהלך הבחינה מוצעים בחוברת הבחינה.

6. הציון בתחום אנגלית במכפ"ל – חלק זה מכיל 54-58 פריטים ומיועד לבחון בקיאות בקריאה והבנה של טקסטים באנגלית ברמה אקדמית. חלק זה מכיל שלושה סוגים של פריטים: השלמת משפטים, ניסוח מחדש והבנת הנקרא. חלק זה משמש בתפקיד כפול: הוא מהווה רכיב של הציון הכללי במכפ"ל והוא גם משמש לשיבוץ הסטודנט בכיתות האנגלית.

סולם עבור כל שלושת תתי-המבחנים: 150-50 עם ממוצע היסטורי של 100 וסטיית תקן 20.

קריטריון

ציון שנה א' (צש"א) – ציון זה הוגדר בחוג כממוצע משוקלל בנקודות זכות (4 לפחות) של ציוני הקורסים אותם למד הסטודנט בשנה א' לתואר בוגר. סולם: 100-40.

3.3. עיבוד הנתונים

כל אחת מהקבוצות 1, 2, 3, 98 ו-99 המהווה קבוצת מוקד (f) הושוותה בנפרד לקבוצה 0 המהווה את קבוצת ההתייחסות (r).
חושבו הסטטיסטיים התיאוריים של משתני המחקר עבור כל אחת מהקבוצות במחקר: ממוצעים, סטיות תקן, והפרש (קריטריון פחות חזאי) של הפערים הנצפים המתוקננים בין ממוצעי קבוצת המוקד לבין קבוצת ההתייחסות (הפרש d -ים של כהן, להלן ה- Δ של כהן) (Thorndike, 1971).
הסטטיסטי התיאורי האחרון מחושב עבור כל אחת משתי שיטות ההשוואה (ראה להלן מחקר מסורתי ומחקר בשיטת ציוני נטייה).

עבור כל אחת משתי שיטות ההשוואה חושבו שני המדדים להגנות:

□ מדד להטיה בביררה על פי המודל של קלירי (להלן, **מדד קלירי** להגנות) – הפער בין ציון הקריטריון הנצפה לבין הציון המנובא על פי קו הרגרסיה המשותף (מחושב במונחים של ציוני תקן).

□ מקדם תוקף (עבור קבוצת ההתייחסות ועבור קבוצת המוקד בנפרד).

הסבר מפורט על המדדים ניתן למצוא בנספחים 7.1.1-7.1.3.

3.3.1. מחקר מסורתי

נכללו בניתוח רק יחידות עיבוד שבהן נמצא לפחות אדם אחד עם "לקות" מקבוצת המוקד המתאימה; ראה תת-הפרק הבא להשוואת שכיחות הלומדים המקורית לזו שנכנסה לניתוח. חושבו הסטטיסטי התיאורי, ה- Δ של כהן, ושני המדדים להגנות, מדד קלירי ומקדם התוקף.

ה- Δ של כהן ומדד קלירי להוגנות חושבו לכל יחידת עיבוד בנפרד ובוצע ממוצע משוקלל בגודל יחידת העיבוד מעבר ליחידות העיבוד. מדד התוקף חושב מעבר ליחידות העיבוד (כי בקבוצות המוקד יש מספר קטן מדי של אנשים לחישוב בתוך יחידת עיבוד)¹⁴.

3.3.1.1. השוואת הקבוצות

להלן לוח המתאר את אוכלוסיית המחקר המקורית וזו שנכנסה לניתוח (לפחות נבחן אחד של קבוצת המוקד בכל יחידת עיבוד).

לוח 4 - שכיחות הלומדים בכל קבוצה בהשוואות השונות

קבוצות	שכיחות מקורית	אחוזים מקורי	שכיחות חדשה	אחוזים חדש
0	120,503	99.21	66,062	98.57
1	958	0.79	958	1.43
0	120,503	99.85	26,225	99.29
2	187	0.15	187	0.71
0	120,503	99.10	71,714	98.49
3	1,096	0.90	1,096	1.51
0	120,503	99.75	35,862	99.17
98	299	0.25	299	0.83
0	120,503	98.80	78,182	98.17
99	1,458	1.20	1,458	1.83

3.3.2. מחקר בשיטת ציוני נטייה

כפי שניתן לראות בלוח 4, בהשוואה מסורתית מושווית קבוצה גדולה בהרבה (קבוצת ההתייחסות) לקבוצה שהיא יחסית קטנה (המוקד). מאחר שהקבוצה הקטנה איננה מדגם מקרי, ההשוואה ביניהן מוטוה. על מנת לערוך השוואה יותר "הוגנת", אנו משתמשים בשיטת ציוני הנטייה כדי לבחור תת-קבוצה מתוך קבוצת ההתייחסות שתהיה שווה בגודלה ודומה יותר במאפייניה לקבוצת המוקד אליה היא תשווה. בשיטה זו לכל אדם מקבוצת המוקד מותאם אדם מקבוצת ההתייחסות שלומד איתו באותה יחידת העיבוד בהתאם למשתני רקע שונים¹⁵. שלא כמו במחקר המסורתי בו העיבודים בוצעו ברמת יחידת העיבוד, בשיטת ציוני הנטייה, בשל מספר הנבדקים הקטן בתוך יחידת העיבוד, החישובים בוצעו מעבר ליחידות העיבוד¹⁶. ראה נספח 7.1.4 להסבר על שיטת ציוני הנטייה.

¹⁴ לכן, בשלב התקנון תיקנו את כל המשתנים (קריטריון וששת החזאים).

¹⁵ ראה תת-הפרק הבא 3.3.2.1 לפירוט.

¹⁶ לכן, בשלב התקנון תיקנו את כל המשתנים (קריטריון וששת החזאים).

3.3.2.1. משתני רקע

לצורך דגימת קבוצות ההשוואה ניסינו בתחילה לזווג לפי חמישה משתני רקע שעליהם נשאלים הנבחנים בעת ההרשמה למכפ"ל: מצב כלכלי, השכלת אם, השכלת אב, מין וגיל. עבור כל אחד ממשתני רקע אלו ועבור כל אחת מהקבוצות נבדקו כמה נבחנים בעלי נתונים יש; בנוסף חושבו ממוצע וסטיית התקן עבור המשתנה. נתונים אלו מוצגים בשני הלוחות הבאים.

לוח 5 - ממוצעים וסטיות תקן עבור משתני הרקע מצב כלכלי¹⁷, השכלת אב והשכלת אם¹⁸

השכלת אם				השכלת אב				מצב כלכלי				קבוצה		
בעלי נתונים	חסרי נתונים	ממוצע	סטיית תקן	בעלי נתונים	חסרי נתונים	ממוצע	סטיית תקן	בעלי נתונים	חסרי נתונים	ממוצע	סטיית תקן	בעלי נתונים	חסרי נתונים	
105,855	14,648	5.20	1.43	105,281	15,222	5.17	1.49	101,099	19,404	3.12	1.04	101,099	19,404	0
207	751	5.33	1.43	207	751	4.98	1.55	190	768	3.24	1.20	190	768	1
60	127	5.65	1.29	62	125	5.48	1.59	58	129	3.09	0.80	58	129	2
271	825	5.00	1.32	270	826	4.92	1.53	265	831	3.21	1.12	265	831	3
124	175	4.73	1.52	124	175	4.66	1.48	115	184	3.46	0.94	115	184	98
462	996	5.19	1.47	459	999	5.18	1.53	426	1,032	3.24	1.09	426	1,032	99
106,979	17,522	5.20	1.43	106,403	18,098	5.17	1.49	102,153	22,348	3.12	1.04	102,153	22,348	כולם

לוח 6 - ממוצעים וסטיות תקן עבור משתני הרקע מין וגיל

גיל				מין			קבוצה	
בעלי נתונים	חסרי נתונים	ממוצע	סטיית תקן	אחוז גברים	אחוז נשים	חסרי נתונים	בעלי נתונים	
119,526	977	21.51	2.49	42	58	9,466	111,037	0
940	18	22.57	2.07	49	51	31	927	1
187	0	22.93	2.29	61	39	5	182	2
1,077	19	22.06	2.41	53	47	35	1,061	3
293	6	21.92	2.36	42	58	9	290	98
1,445	13	22.04	2.06	42	58	49	1,409	99
123,468	1,033	21.53	2.48	42	58	9,595	114,906	כולם

¹⁷ סולם: 1 = גבוה בהרבה מהממוצע... 6 = נמוך בהרבה מהממוצע.

¹⁸ סולם: 1 = לא למד כלל 2 = יסודית 3 = תיכונית חלקית 4 = תיכונית מלאה 5 = על-תיכונית או אקדמית חלקית 6 = תואר ראשון 7 = תואר שני או שלישי.

כפי שניתן לראות בלוח 5 עבור שלושת משתני הרקע מצב כלכלי, השכלת אם והשכלת אב, יש בקבוצות המוקד הרבה יותר אנשים עם נתונים חסרים מאשר כאלה בעלי נתונים. לעומת זאת, לוח 6 מראה שעבור מין וגיל המצב הרבה יותר טוב וכמעט שאין אנשים עם נתונים חסרים.

3.3.2.2. השלמת נתונים

לאור הממצאים בלוח 5 ובלוח 6 היה ניסיון להשלים נתונים חסרים במשתנים מצב כלכלי, השכלת אם והשכלת אב מהיבחנויות קודמות. השלמת הנתונים בוצעה בהתאם לכלל הבא:

□ אם בבחינה הקובעת המשתנה חסר, הסתכלנו על בחינות נוספות של אותו אדם (לצורך זה קובץ התוקף הוצלב עם הקובץ הפסיכומטרי לפי מספר תעודת זהות); הנתון הושלם במידה ולא קיימות סתירות בדיווחים האחרים.

להלן לוח המתאר את התפלגות הקבוצות השונות לאחר השלמת הנתונים עבור המשתנים מצב כלכלי, השכלת אם והשכלת אב.

לוח 7 – ממוצעים וסטיות תקן עבור משתני הרקע מצב כלכלי, השכלת אב והשכלת אם לאחר השלמת נתונים חסרים

השכלת אם				השכלת אב				מצב כלכלי				קבוצה
ממוצע	אחוז נתונים	חסרי נתונים	בעלי נתונים	ממוצע	אחוז נתונים	חסרי נתונים	בעלי נתונים	ממוצע	אחוז נתונים	חסרי נתונים	בעלי נתונים	
5.19	90	12,036	108,467	5.17	90	12,576	107,927	3.12	86	16,476	104,027	0
5.26	30	670	288	4.98	30	672	286	3.23	28	690	268	1
5.57	40	113	74	5.59	40	112	75	2.99	39	115	72	2
5.09	38	675	421	5.05	38	675	421	3.11	37	688	408	3
4.81	51	148	151	4.75	50	149	150	3.47	45	164	135	98
5.23	46	792	666	5.22	45	804	654	3.26	42	847	611	99

כפי שניתן לראות מהסתכלות בלוח 5 ובלוח 7 המצב אכן השתפר, אבל עדיין חסרים מספר רב של נתונים (יותר חסרים מקיימים). לכן, לא יהיה זה נכון להוציא מן המחקר אנשים עם נתונים חסרים. אפשרות נוספת היא לנסות ולהשלים לכל אדם שחסרים לגביו נתונים את ממוצע הקבוצה אליה הוא משתייך. אי אפשר להניח שהממוצעים שמיוצגים על ידי בעלי הנתונים מייצגים את כלל הקבוצה.

בהתאם לכך, הוחלט לעשות את הזיווג על פי מין וגיל בלבד (כאשר לא כלולים במחקר אנשים בעלי נתונים חסרים במשתנים אלו) ולוותר על שלושת המשתנים האלה כמשתני רקע.

3.3.2.3. זיווג הקבוצות

השוואת הקבוצות בוצעה בשני שלבים: קודם כל (בדומה להשוואה המסורתית) נבחרו רק יחידות עיבוד עבורן נמצא לפחות נבחן אחד עם "לקות" מקבוצת המוקד המתאימה. בנוסף, בשלב זה גם הוסרו בעלי נתונים חסרים. בשלב השני הופעל אלגוריתם הזיווג אשר לאחר מיון אקראי של הנבחים, מתאים לכל נבחן במוקד בן-זוג בקבוצת ההתייחסות מאותה יחידת עיבוד שלו ציון נטייה הקרוב ביותר לשלו.

בלוח 8 מוצגת שכיחות והתפלגות התצפיות על פי משתנה הרקע לפני ואחרי בחירת תצפיות מקבוצת ההתייחסות על ידי השוואת ציוני הנטייה.

לוח 8 - גודל ההשוואה עם יחידות העיבוד הרלוונטיות לפני ואחרי הזיווג

לאחר ההשוואה לפי ציון נטייה					לפני ההשוואה לפי ציון נטייה					קבוצות
גיל		מין	אנשים		גיל		מין	אנשים		
ממוצע	סטיית תקן	אחוז גברים	אחוז נשים		ממוצע	סטיית תקן	אחוז גברים	אחוז נשים		
22.51	2.09	49	51	913	21.62	2.53	40	60	59,283	0
22.58	2.07	49	51	913	22.58	2.07	49	51	913	1
22.86	2.16	61	39	182	21.64	2.59	40	60	23,571	0
22.91	2.32	61	39	182	22.91	2.32	61	39	182	2
21.96	2.07	53	47	1,046	21.51	2.51	41	59	64,930	0
22.02	2.32	53	47	1,046	22.02	2.32	53	47	1,046	3
21.91	2.59	43	57	286	21.63	2.67	37	63	32,586	0
21.92	2.36	42	58	286	21.92	2.36	42	58	286	98
22.03	2.08	43	57	1,398	21.56	2.49	41	59	70,615	0
22.06	2.05	43	57	1,398	22.06	2.05	43	57	1,398	99

* מוצגים רק בעלי נתונים מלאים במשתנים מין וגיל.

** מוצגים נבחים מיחידות עיבוד שבהן לפחות נבחן אחד עם "לקות" במוקד.

כפי שניתן לראות בלוח 8 מידת הדמיון, הן מבחינת התפלגות המין והן מבחינת הגיל הממוצע, בין הקבוצות גדלה בעקבות הזיווג. ניתן לומר שהשיטה הצליחה לזווג בצורה טובה בין שתי הקבוצות לגבי משתנים אלו.

4. תוצאות

ממוצעים וסטיות תקן עבור הקריטריון וששת החזאים עבור כל אחת משש הקבוצות מוצגים בלוח הבא. החישובים מתייחסים למערך המקורי שמוצג בלוח 3.

לוח 9 - ממוצע וסטיית תקן עבור הקריטריון והחזאים
משתנה

משתנה														קבוצה
אנגלית		כמותי		מילולי		סכמ		מכפ"ל		בגרות		צ"ש"א		
ממוצע	סטיית	ממוצע	סטיית	ממוצע	סטיית	ממוצע	סטיית	ממוצע	סטיית	ממוצע	סטיית	ממוצע	סטיית	
תקן	תקן	תקן	תקן	תקן	תקן	תקן	תקן	תקן	תקן	תקן	תקן	תקן	תקן	
124.6	17.0	122.6	16.4	122.4	9.4	59.1	85.0	630.5	8.8	99.0	9.6	82.2	0	
114.7	20.4	116.6	15.7	116.8	8.7	54.1	83.8	593.2	7.9	94.2	9.8	80.7	1	
120.7	18.1	119.5	14.8	120.1	8.5	55.2	83.0	613.7	7.3	94.3	10.6	80.8	2	
122.9	17.6	122.9	16.1	123.6	9.4	58.8	87.0	632.2	8.3	98.6	10.7	80.0	3	
109.1	19.7	109.4	18.3	107.0	10.2	51.4	97.3	547.9	8.9	94.3	9.3	81.8	98	
115.3	17.7	115.5	16.7	114.4	9.4	54.6	87.6	585.9	8.6	96.1	9.8	81.8	99	

בהתבסס על הנתונים בלוח 9 חושבו מדדי ה- Δ של כהן¹⁹. כל התוצאות הן יחסית לקבוצת ההתייחסות המתאימה (ערך חיובי הוא לטובת קבוצת המוקד²⁰). התוצאות מוצגות עבור שתי שיטות ההשוואה: מסורתית (לוח 10 ולוח 11) וזו שבעזרת ציוני הנטייה (לוח 12 ולוח 13).

לוח 10 - ה- Δ של כהן עבור חזאים מרכזיים - מחקר מסורתי

חזאי									קבוצה
סכמ			מכפ"ל			בגרות			
$\Delta = d_{\text{חזאי}} - d_{\text{קריטריון}}$	$d_{\text{קריטריון}}$	$d_{\text{חזאי}}$	$\Delta = d_{\text{חזאי}} - d_{\text{קריטריון}}$	$d_{\text{קריטריון}}$	$d_{\text{חזאי}}$	$\Delta = d_{\text{חזאי}} - d_{\text{קריטריון}}$	$d_{\text{קריטריון}}$	$d_{\text{חזאי}}$	
-0.18	-0.19	-0.37	-0.03	-0.19	-0.22	-0.17	-0.19	-0.36	1
-0.21	-0.08	-0.29	0.08	-0.08	0.00	-0.35	-0.08	-0.43	2
0.26	-0.36	-0.10	0.32	-0.36	-0.04	0.26	-0.36	-0.10	3
-0.49	-0.06	-0.55	-0.66	-0.06	-0.72	-0.16	-0.06	-0.22	98
-0.29	-0.13	-0.42	-0.36	-0.13	-0.49	-0.06	-0.13	-0.19	99

¹⁹ הערך המוצג מתוקן לחוסר מהימנות הקריטריון בהתאם לנוסחה: $r'_{xy} = \frac{r_{xy}}{\sqrt{r_{xx}} \cdot \sqrt{r_{yy}}}$ (Lord & Novick, 1968)
 נוסחה 3.9.6 עמוד 70) שמוכנסת לנוסחה השנייה בנספח 7.1.1. הערך שנלקח לתיקון המהימנות הוא $r_{yy} = 0.68$ (קנת-כהן ואחרים, 1999).
²⁰ ראה עמוד 14 או נספח 7.1.1 להסבר מדויק על הכיוונים של החישובים.

לוח 11 - ה- Δ של כהן עבור חזאים משניים - מחקר מסורתי

חזאי									קבוצה
אנגלית			כמותי			מילולי			
$\Delta = d_{\text{חזאי}} - d_{\text{קריטריון}}$	$d_{\text{קריטריון}}$	$d_{\text{חזאי}}$	$\Delta = d_{\text{חזאי}} - d_{\text{קריטריון}}$	$d_{\text{קריטריון}}$	$d_{\text{חזאי}}$	$\Delta = d_{\text{חזאי}} - d_{\text{קריטריון}}$	$d_{\text{קריטריון}}$	$d_{\text{חזאי}}$	
-0.25	-0.19	-0.44	0.12	-0.19	-0.07	0.09	-0.19	-0.10	1
0.06	-0.08	-0.02	0.06	-0.08	-0.02	0.16	-0.08	0.08	2
0.21	-0.36	-0.15	0.35	-0.36	-0.01	0.39	-0.36	0.03	3
-0.44	-0.06	-0.50	-0.46	-0.06	-0.52	-0.64	-0.06	-0.70	98
-0.30	-0.13	-0.43	-0.18	-0.13	-0.31	-0.29	-0.13	-0.42	99

לוח 12 - ה- Δ של כהן עבור חזאים מרכזיים - מחקר ציוני נטייה

חזאי									קבוצה
סכמ			מכפ"ל			בגרות			
$\Delta = d_{\text{חזאי}} - d_{\text{קריטריון}}$	$d_{\text{קריטריון}}$	$d_{\text{חזאי}}$	$\Delta = d_{\text{חזאי}} - d_{\text{קריטריון}}$	$d_{\text{קריטריון}}$	$d_{\text{חזאי}}$	$\Delta = d_{\text{חזאי}} - d_{\text{קריטריון}}$	$d_{\text{קריטריון}}$	$d_{\text{חזאי}}$	
-0.17	-0.16	-0.33	-0.17	-0.16	-0.33	-0.05	-0.16	-0.21	1
-0.13	-0.12	-0.25	-0.14	-0.12	-0.26	-0.07	-0.12	-0.19	2
0.25	-0.37	-0.12	0.23	-0.37	-0.14	0.32	-0.37	-0.05	3
-0.50	0.05	-0.45	-0.71	0.05	-0.66	-0.13	0.05	-0.08	98
-0.29	-0.12	-0.41	-0.41	-0.12	-0.53	0.00	-0.12	-0.12	99

לוח 13 - ה- Δ של כהן עבור חזאים משניים - מחקר ציוני נטייה

חזאי									קבוצה
אנגלית			כמותי			מילולי			
$\Delta = d_{\text{חזאי}} - d_{\text{קריטריון}}$	$d_{\text{קריטריון}}$	$d_{\text{חזאי}}$	$\Delta = d_{\text{חזאי}} - d_{\text{קריטריון}}$	$d_{\text{קריטריון}}$	$d_{\text{חזאי}}$	$\Delta = d_{\text{חזאי}} - d_{\text{קריטריון}}$	$d_{\text{קריטריון}}$	$d_{\text{חזאי}}$	
-0.29	-0.16	-0.45	0.05	-0.16	-0.11	-0.07	-0.16	-0.23	1
-0.07	-0.12	-0.19	-0.08	-0.12	-0.20	-0.05	-0.12	-0.17	2
0.14	-0.37	-0.23	0.24	-0.37	-0.13	0.35	-0.37	-0.02	3
-0.58	0.05	-0.53	-0.48	0.05	-0.43	-0.62	0.05	-0.57	98
-0.34	-0.12	-0.46	-0.22	-0.12	-0.34	-0.32	-0.12	-0.44	99

4.1. תוצאות מפורטות

בחלק זה מוצגים בטבלאות מדד קלירי להטיה בברירה ומקדם התוקף בהתאם לשתי שיטות המחקר. בפרק 5.3 מוצגות אותן התוצאות בצורה גרפית.

4.1.1. מחקר מסורתי

מוצגים שני לוחות – מדד קלירי להטיה ומקדם התוקף.

לוח 14 - מדדי הטיה* ביחס לקבוצת ההתייחסות על פי קלירי - מחקר מסורתי

	קבוצה					
	בגרות	מכפ"ל	סכמ	מילולי	כמותי	אנגלית
1	0.03	0.08	0.01	0.12	0.12	0.11
2	-0.07	0.07	-0.07	0.11	0.08	0.06
3	0.24	0.26	0.23	0.28	0.27	0.26
98	-0.03	-0.16	-0.16	-0.14	-0.09	-0.03
99	0.02	-0.05	-0.07	-0.01	-0.01	0.02

*במונחי סטיות תקן של הקריטריון. ערך חיובי מסמן הטיה לטובת קבוצת המוקד וערך שלילי מסמן הטיה לרעתה.

לוח 15 - מקדם תוקף* - מחקר מסורתי

	חזאי										קבוצה	
	בגרות		מכפ"ל		סכמ		מילולי		כמותי			אנגלית
	מוקד	התייחסות	מוקד	התייחסות	מוקד	התייחסות	מוקד	התייחסות	מוקד	התייחסות	מוקד	התייחסות
1	0.23	0.35	0.31	0.28	0.33	0.40	0.28	0.28	0.24	0.27	0.26	0.16
2	0.26	0.36	0.28	0.29	0.34	0.40	0.29	0.26	0.14	0.27	0.26	0.17
3	0.29	0.34	0.28	0.26	0.37	0.38	0.26	0.23	0.22	0.30	0.25	0.13
98	0.33	0.35	0.28	0.29	0.37	0.38	0.29	0.26	0.27	0.24	0.27	0.17
99	0.29	0.34	0.24	0.27	0.35	0.39	0.27	0.22	0.23	0.24	0.26	0.14

*לא בוצע תיקון לקיצוץ תחום

4.1.2. מחקר בשיטת ציוני נטייה

מוצגים שני לוחות – מדד קלירי להטיה ומקדם התוקף.

לוח 16 - מדדי הטיה* ביחס לקבוצת ההתייחסות על פי קלירי - מחקר ציוני נטייה

	קבוצה					
	חזאי					
	אנגלית	כמותי	מילולי	סכמ	מכפ"ל	בגרות
1	0.04	0.05	0.04	0.01	0.02	0.03
2	0.03	0.02	0.03	0.00	0.01	0.02
3	0.14	0.14	0.14	0.14	0.14	0.15
98	-0.05	-0.06	-0.08	-0.09	-0.10	-0.03
99	0.02	0.01	0.00	-0.02	-0.02	0.03

*במונחי סטיות תקן של הקריטריון. ערך חיובי מסמן הטיה לטובת קבוצת המוקד וערך שלילי מסמן הטיה לרעתה.

לוח 17 - מקדם תוקף* - מחקר ציוני נטייה

	קבוצה											
	חזאי											
	אנגלית		כמותי		מילולי		סכמ		מכפ"ל		בגרות	
	התייחסות	מוקד	התייחסות	מוקד	התייחסות	מוקד	התייחסות	מוקד	התייחסות	מוקד	התייחסות	מוקד
1	0.20	0.08	0.24	0.25	0.23	0.28	0.40	0.33	0.27	0.30	0.38	0.23
2	0.12	0.27	0.36	0.26	0.29	0.14	0.48	0.33	0.33	0.29	0.43	0.24
3	0.14	0.12	0.21	0.28	0.18	0.23	0.37	0.37	0.23	0.27	0.34	0.30
98	0.15	0.12	0.23	0.23	0.28	0.25	0.34	0.36	0.28	0.26	0.27	0.33
99	0.15	0.09	0.25	0.24	0.20	0.23	0.37	0.35	0.25	0.25	0.33	0.28

*לא בוצע תיקון לקיצוץ תחום

4.2. סיכום תוצאות

לצורך דיון בעקביות התוצאות יהיה יותר קל להציג את התוצאות בלוח 10 עד לוח 14 ובלוח 16 בצורה ממצה יותר. התוצאות מוצגות בהתאם לכלל האצבע של כהן (1998) לגודל האפקט. ראה מקרא להסבר הסימונים ב-לוח 20.

לוח 18 - סיכום תוצאות: חזאים מרכזיים

חזאי												קבוצה
סכמ				מכפ"ל				בגרות				
CP	CT	DP	DT	CP	CT	DP	DT	CP	CT	DP	DT	
		-	-			-						1 = לקויות למידה
			-									2 = קשב וריכוז
	+	+	+		+	+	+		+	+	+	3 = מגבלות פיזיות
	-	--	--		-	---	---					98 = "חסרי נתונים"
		-	-			-	-					99 = מסורבים מקצועית

לוח 19 - סיכום תוצאות: חזאים משניים

חזאי												קבוצה
אנגלית				כמותי				מילולי				
CP	CT	DP	DT	CP	CT	DP	DT	CP	CT	DP	DT	
		-	-									1 = לקויות למידה
											+	2 = קשב וריכוז
	+		+		+	+	+		+	+	+	3 = מגבלות פיזיות
		--	-			--	--			--	-	98 = "חסרי נתונים"
		-	-			-	-			-	-	99 = מסורבים מקצועית

לוח 20 - מקרא

סימון	מדד	סימון	גודל אפקט	טווח
DT	כהן מסורתי		זניח	0.00-0.14
DP	כהן נטייה	+/-	קטן	0.15-0.44
CT	קלירי מסורתי	++/--	בינוני	0.45-0.64
CP	קלירי נטייה	+++/-	גדול	0.65-

ננתח פה רק את נושא העקביות בין השיטות ובתוך המדדים השונים. דיון נוסף בנושא הבדלים בין השיטה המסורתית לשיטת ציוני הנטייה ניתן למצוא בתת-פרק 5.3 ובנספח 7.2.

דיון בתוצאות עצמן נמצא בפרק הבא (פרק 5).

ניתן לראות בלוח 18 ובלוח 19 כי עבור כל אחת מחמש קבוצות המוקד ועבור חזאי מסוים אין סתירות בתוצאות, במובן שאין סימנים מנוגדים באותה השורה (עובדה שנכונה גם מעבר לחזאי בודד). בנוסף, בתוך אותו מדד (כהן או קלירי) ועבור חזאי מסוים אין הבדל של יותר מדרגה אחת בין מחקר מסורתי ומחקר בשיטת ציוני נטייה.

בנוסף, הצגה זו של התוצאות מראה בבירור כיצד הערך בסכם (שהוא כאמור ממוצע במשקולות שווים של מכפ"ל והבגרות) נבנה. כך למשל, שני המינוסים ב-DT בקבוצה 98 עבור החזאי סכם נוצרים משלושה מינוסים ומינוס אחד בחזאים מכפ"ל והבגרות בהתאמה. לבסוף, ניתן לראות כיצד תוצאת מכפ"ל מתקבלת משלושת רכיביה: מילולי, כמותי ואנגלית. כך למשל, עבור קבוצה 99 המינוס היחיד ב-DP נוצר ממינוס יחיד בשלושת הרכיבים בעוד שעבור קבוצה 1 המינוס היחיד במדד זה נוצר רק על ידי מינוס יחיד ברכיב האנגלית.

עוד ניתן לראות בלוח 18 שעבור קבוצות 1 ו-2 יש לכאורה הבדלים בין השיטה המסורתית ושיטת ציוני הנטייה לגבי ה- Δ של כהן עבור הבגרות ומכפ"ל. ניתוח מעמיק יותר לתופעה זו ניתן למצוא בנספח 7.2.

5.1. הטיה בברירה

המדדים כהן וקלירי הם אומנם לכאורה בני-השוואה (שניהם במונחי סטיות התקן של הקריטריון) אבל אין זה אומר שהם אמורים להיות שווים – כל אחד יוצא מנקודת מוצא שונה לגבי משמעות המושג הוגנות.

נקודת המוצא של ה- Δ של כהן היא שברירה הכרחית (שבה הסיבה היחידה לדחיית מועמדים היא שמספרם גדול ממספר המקומות – המצב השכיח בהקשר של ברירה להשכלה גבוהה) 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.

לעומתו, קלירי יוצא מתוך נקודת הנחה מריטוקרטית להוגנות שטוענת שהוגן לקבל את אלו בעלי הציון הגבוה ביותר (ראה (Darlington, 1971).

הביקורת העיקרית כנגד מדד ה- Δ של כהן כמדד להוגנות (נבון וכהן, 2009) היא שהוא מתעלם לחלוטין 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 הברירה.

מנגד, בסדרה של עבודות הראו לין ועמיתיו (Linn, 1983a, 1983b; Linn & Hastings, 1984; Linn & Wert, 1971) שהממצא הרווח של הטיה לטובת קבוצות חלשות במחקרים המשתמשים במודל הרגרסיה של קלירי נובע מארטיפקטים סטטיסטיים. שיטה חלופית לאיתור הטיה בברירה, ברוח המודל של קלירי (בדיקת קיומו של ניבוי דיפרנציאלי), שמתמודדת עם הבעייתיות הזו (Linn, 1984) יושמה במחקרי הוגנות קודמים שנעשו במאליו (Turvall, Bronner, Kennet-Cohen & Oren, 2008; קנת-כהן, טורוול ואורן, 2011) שיטה זו אינה ישימה בהקשר הנוכחי בשל גודל המדגם. לאור האמור לעיל מומלץ להתייחס בזהירות לממצאים המעידים על הטיה לטובת הקבוצות החלשות. לעומת זאת, ממצאים על הטיה לרעת הקבוצות החלשות אינם טריוויאליים וראויים להתייחסות.

מאחר שבמקרה שלנו אין הממצאים העיקריים מראים הטיה לטובת קבוצות המוקד ולפיכך אין התוצאות חשודות כארטיפקטואליות (ברוח הביקורת של לין ועמיתיו שלעיל) וכן בהתאמה עם קונצנזוס רחב בקרב חוקרי ההטיה שמפורסמים בסטנדרטים החשובים לבניית מבחנים (AERA/APA/NCME, 1999), נשתמש במדדי קלירי בנייתוח ההטיה בברירה.

הניתוח הבא של ההטיה בברירה מתייחס ללוח 18 ולוח 19 ואיור 6 עד איור 10.

- לקויות למידה, קשב וריכוז, מסורבים מקצועית – קלירי מראה שאין הטיה בכל החזאים (נכון לגבי שתי השיטות).
- מגבלות פיזיות – עבור קבוצה זו שתי השיטות מראות ניבוי יתר קל עבור כל החזאים.
קבוצה זו של אנשים בעלי מגבלות פיזיות כוללת בתוכה גם אנשים עם בעיות סכרת, הזעת יתר וכו', בעיות שלא בהכרח באות לידי ביטוי במבחן עצמו. עבור אנשים אלו מתן ההקלות מקל כנראה יתר על המידה ומכאן ניבוי היתר הקל. בכל מקרה, עבור קבוצה זו של אנשים הכוללת, בין השאר, אנשים עם נכויות קשות המערכת תעדיף את המצב הקיים של ההטיה הקלה לטובתם.
- "חסרי נתונים" – מדד קלירי מסורתי מראה הטיה קטנה לרעתם במכפ"ל וכתוצאה מכך בסכם (לפי שיטת ציוני הנטייה האפקט קטן יותר ולפי כלל האצבע של כהן לגודל ההטיה לא נמצאה שום הטיה עבור קבוצה זו).
נעיר שעבור קבוצה זו של "חסרי נתונים" יתכן מאוד וחלק גדול מהאנשים, לו היו ממצאים את כל המסמכים בזמן, כנראה היו מקבלים תנאים; לכן ממצא זה הוא צפוי.
לבסוף, נוסיף ונאמר, שעבור שתי הקבוצות האחרונות, "חסרי הנתונים" והמסורבים מקצועית, אין לנו מידע לגבי ההתאמות שאנשים אלו קיבלו בלימודים באוניברסיטה²¹. רק בדיקה מעמיקה יותר הלוקחת בחשבון את התנאים האלו יכולה לשפוך אור מדויק יותר לגבי ההוגנות האמיתית של מכפ"ל.

²¹ יתכן שתנאים מותאמים במבחנים באוניברסיטה ניתנים יחסית ביתר קלות.

5.2. תוקף דיפרנציאלי

כאמור, מכיוון שבכל יחידת עיבוד יכולנו למצוא מספר קטן מאוד של סטודנטים מקבוצות המוקד (לפעמים רק אחד), נאלצנו לחשב את התוקף מעבר ליחידות העיבוד השונות. מגבלה זו מונעת מאיתנו לבצע תיקון לקיצוץ תחום שבא לתקן ארטיפקט סטטיסטי. לכן, את מקדמי התוקף המוצגים בלוח 15 ובלוח 17 יש לבחון בצורה השוואתית בלבד ביחס לקבוצות ההתייחסות המתאימות אשר לגביהן חושב התוקף באותם התנאים.

כפי שניתן לראות באיור 11 עד איור 15 אשר בפרק 5.3 התוצאות עבור המחקר המסורתי והמחקר בשיטת ציוני הנטייה דומות ולכן לא נתייחס אליהן בנפרד.

באופן כללי ניתן לומר, שכפי שמצופה, התוקף של ציון הסכם גבוה יותר מזה של הבררות ומכפ"ל בנפרד. זה האחרון, התוקף של מכפ"ל, גדול יותר משלושת רכיביו: מילולי, כמותי ואנגלית. מסקנה זו נכונה עבור כל חמש קבוצות המוקד וקבוצת ההתייחסות המתאימה לכל אחת בנפרד. משמעות תוצאה זו היא ששני הציונים הקובעים לקבלה לאוניברסיטה, הבררות וציון מכפ"ל, מכילים ערך מוסף לזה של כל אחד בנפרד.

השוואת התוקף של קבוצת המוקד לזה של קבוצת ההתייחסות מראה שתוקף הניבוי של הבררות וציון הסכם גבוה יותר עבור קבוצת ההתייחסות (פרט לקבוצת "חסרי הנתונים"). תופעה זו בולטת במיוחד בקבוצת לקויי הלמידה ובקבוצת בעיות קשב וריכוז. עבור מכפ"ל התוקף של קבוצת המוקד וההתייחסות די דומה. ביחס למרכיבי מכפ"ל התמונה מעורבת; עבור חלק מהקבוצות והחזאים (לקויי למידה ומגבלות פיזיות בחלק המילולי והכמותי) התוקף גבוה יותר עבור קבוצת המוקד, בעוד שעבור האחרים (בעיות קשב וריכוז בחלק המילולי והכמותי) קיים יתרון קל עבור קבוצת ההתייחסות. לגבי שתי הקבוצות האחרונות, "חסרי הנתונים" ומסורבים מקצועית, החלקים המילוליים והכמותיים דומים מבחינת התוקף של קבוצת המוקד וההתייחסות.

לגבי החלק באנגלית, ניכר שחלק זה הינו בעל תוקף הניבוי הנמוך ביותר יחסית לשאר החזאים (פרט לקבוצת בעיות קשב וריכוז). תוצאה זו נכונה עבור קבוצות המוקד ועבור קבוצות ההתייחסות, כאשר בקבוצות לקויי הלמידה והמסורבים מקצועית התוקף בקבוצת ההתייחסות גבוה יותר, עבור קבוצת בעיות קשב וריכוז התוקף במוקד גבוה יותר ובשאר שתי הקבוצות, מגבלות פיזיות ו"חסרי הנתונים", התוקף דומה בקבוצת המוקד ובקבוצת ההתייחסות.

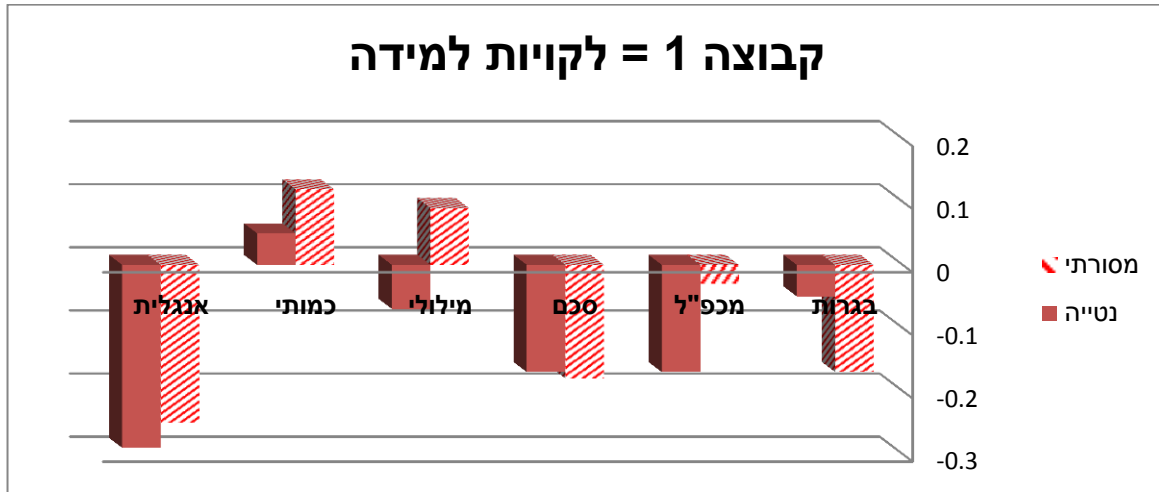
באופן כללי ניתן לומר שדפוס התוצאות שהתקבל במחקר זה הן לגבי ההטיה והן לגבי התוקף דומה לממצאים המדווחים במחקר קודם שנערך במאלי"ו (Oren & Even, 2005).

5.3. השוואה בין השיטות

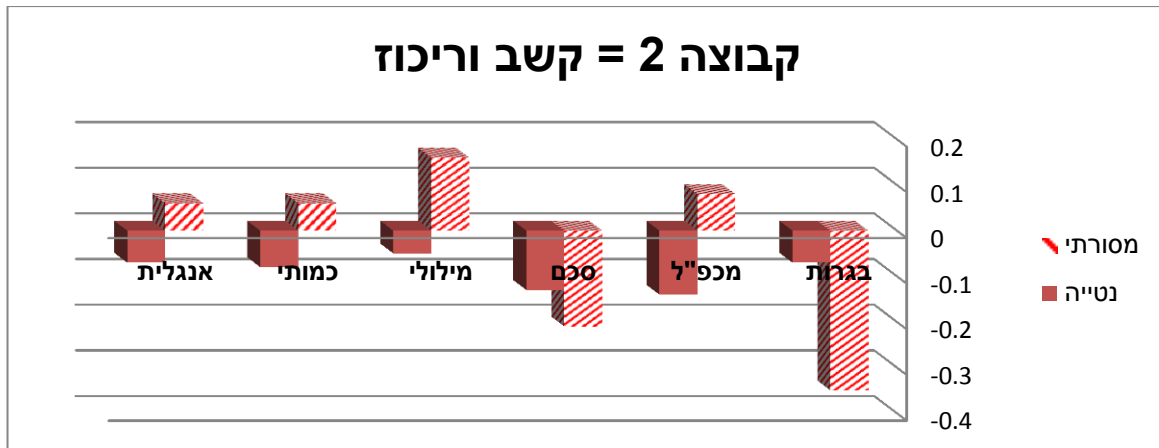
לצורך השוואה בין שתי שיטות המחקר, המסורתית ושיטת ציוני הנטייה, נציג את התוצאות של פרק 4 בצורה גרפית.

5.3.1. ה- Δ של כהן

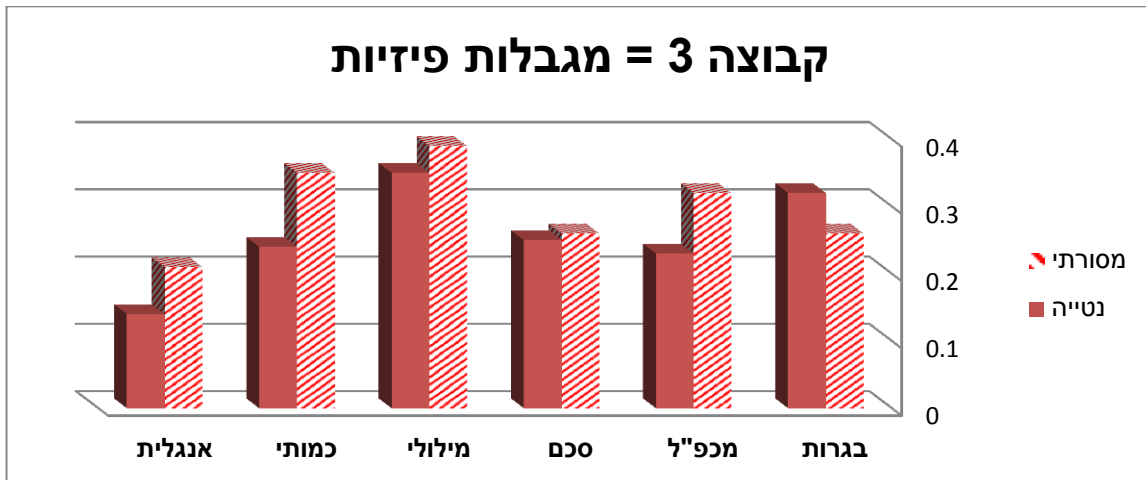
איור 1 - ה- Δ של כהן עבור קבוצה 1 = לקויות למידה



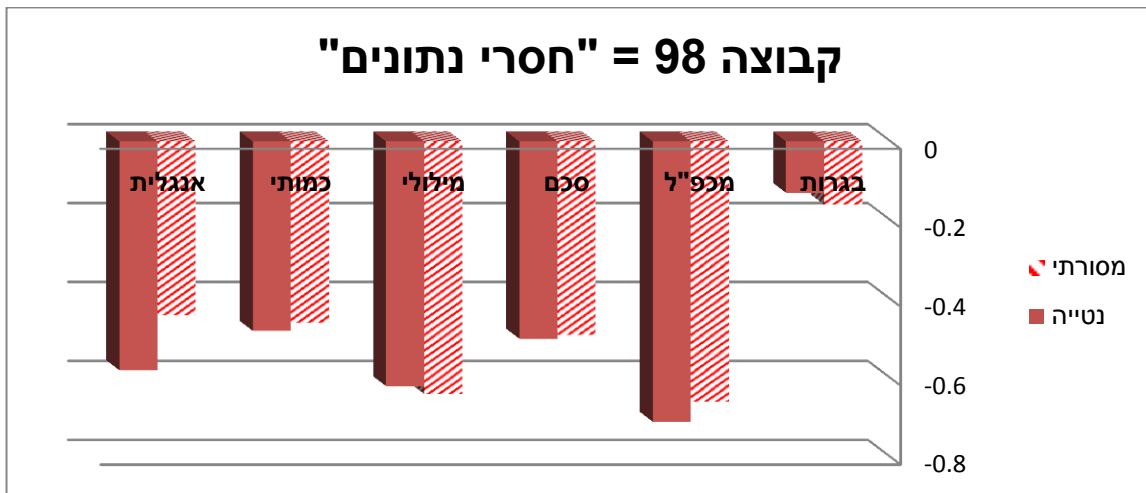
איור 2 - ה- Δ של כהן עבור קבוצה 2 = קשב וריכוז



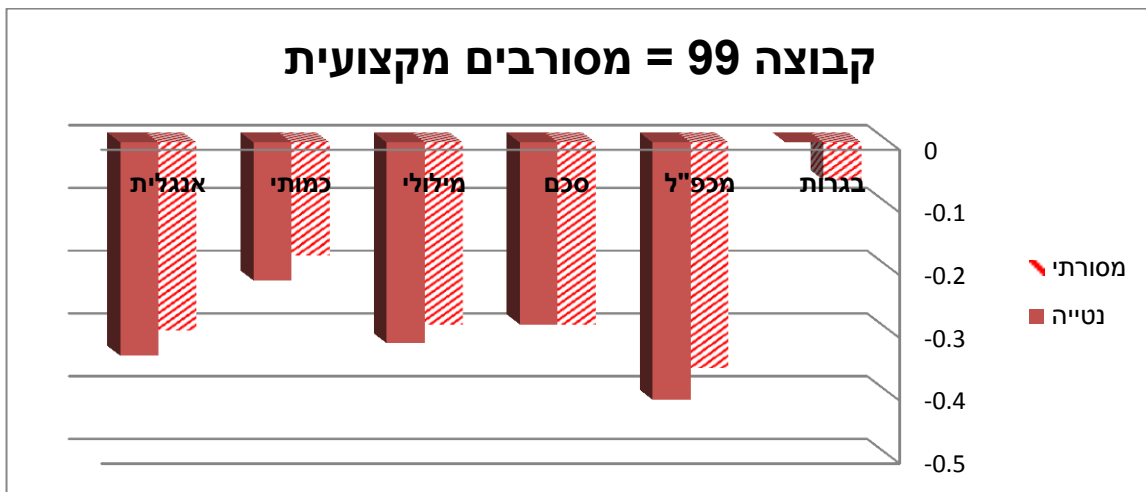
איור 3 - ה- Δ של כהן עבור קבוצה 3 = מגבלות פיזיות



איור 4 - ה- Δ של כהן עבור קבוצה 98 = "חסרי נתונים"



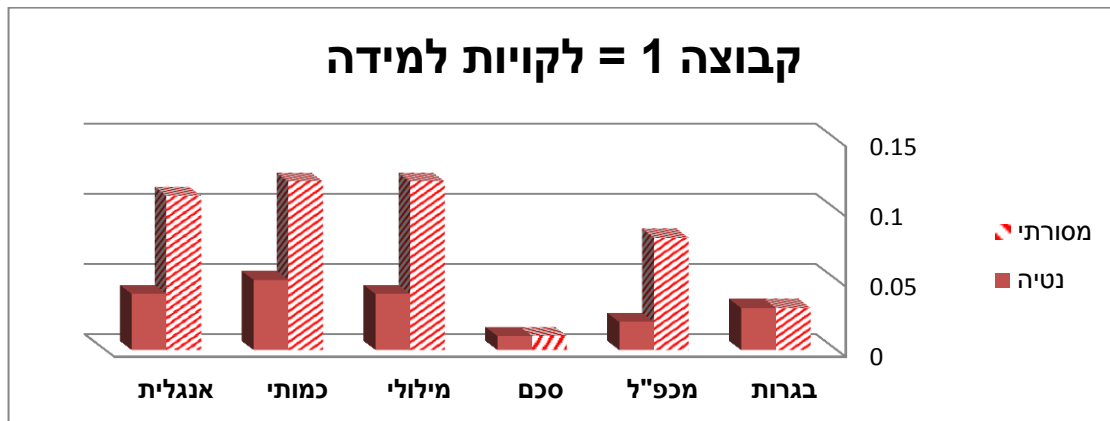
איור 5 - ה- Δ של כהן עבור קבוצה 99 = מסורבים מקצועית



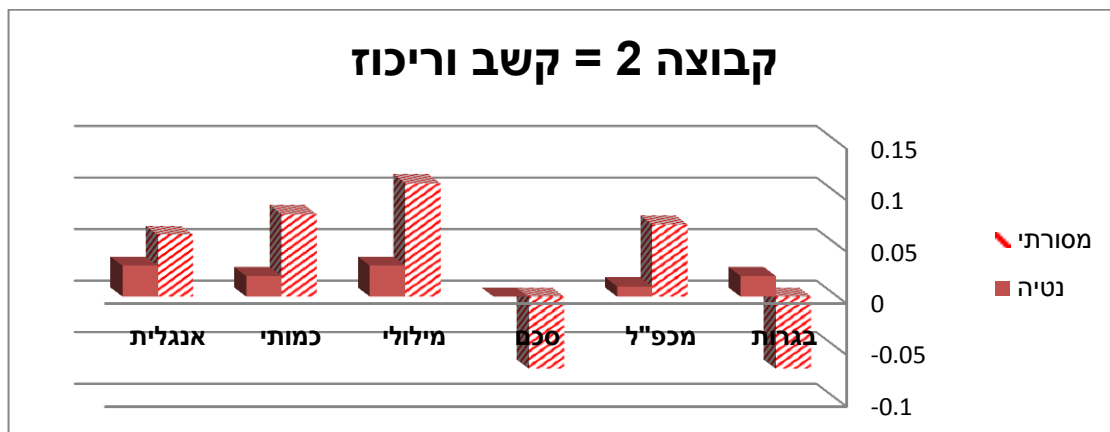
מהתבוננות באיור 1 עד איור 5 ניתן לראות שעבור קבוצות 98,3 ו-99 שתי השיטות נותנות בערך אותו ערך של ה- Δ של כהן. עבור קבוצות 1 ו-2 יש הבדלים בבגרות ובמכפ"ל (וכתוצאה מכך בחלק מתתי-המבחנים); הבדלים אלו לפעמים הם אף הבדלים בסימן. ניתוח מעמיק לתופעה זו ניתן למצוא בנספח 7.2.

5.3.2. מדד קלירי להגנות

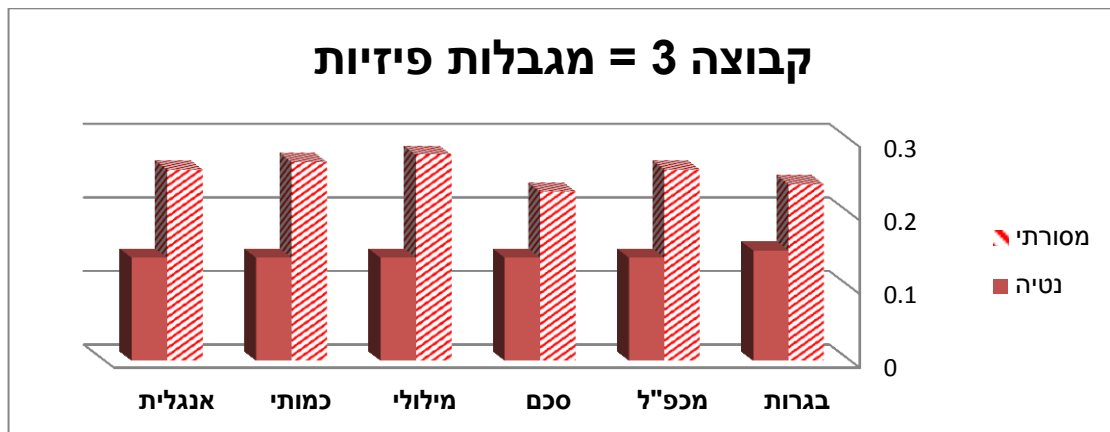
איור 6 - מדד קלירי עבור קבוצה 1 = לקויות למידה



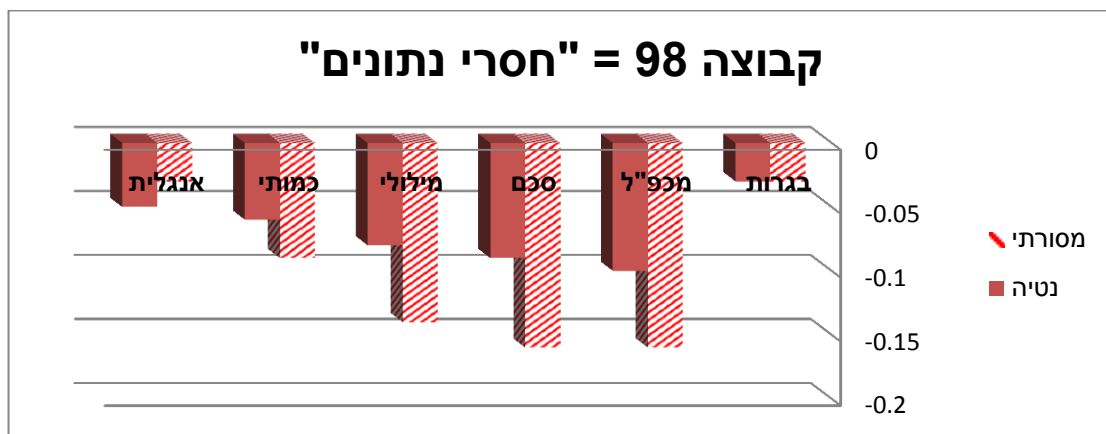
איור 7 - מדד קלירי עבור קבוצה 2 = קשב וריכוז



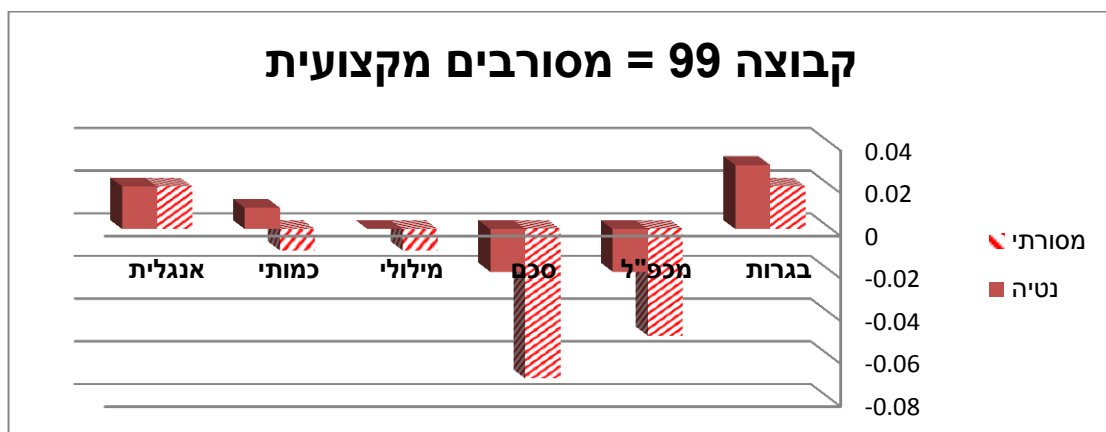
איור 8 - מדד קלירי עבור קבוצה 3 = מגבלות פיזיות



איור 9 - מדד קלירי עבור קבוצה 98 = "חסרי נתונים"



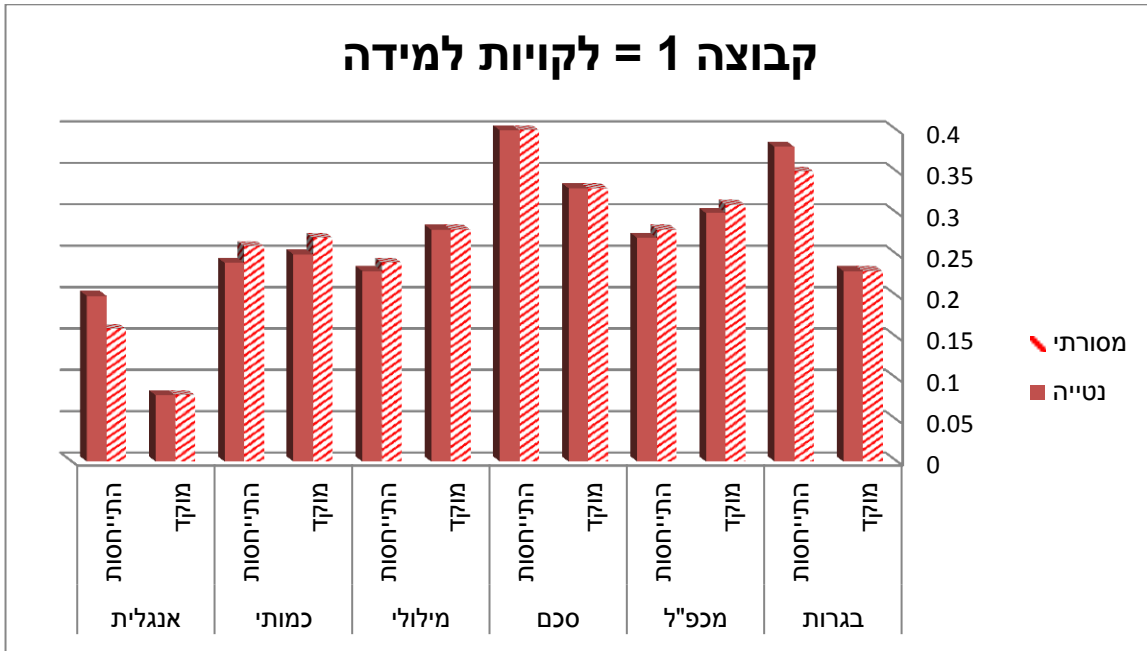
איור 10 - מדד קלירי עבור קבוצה 99 = מסורבים מקצועית



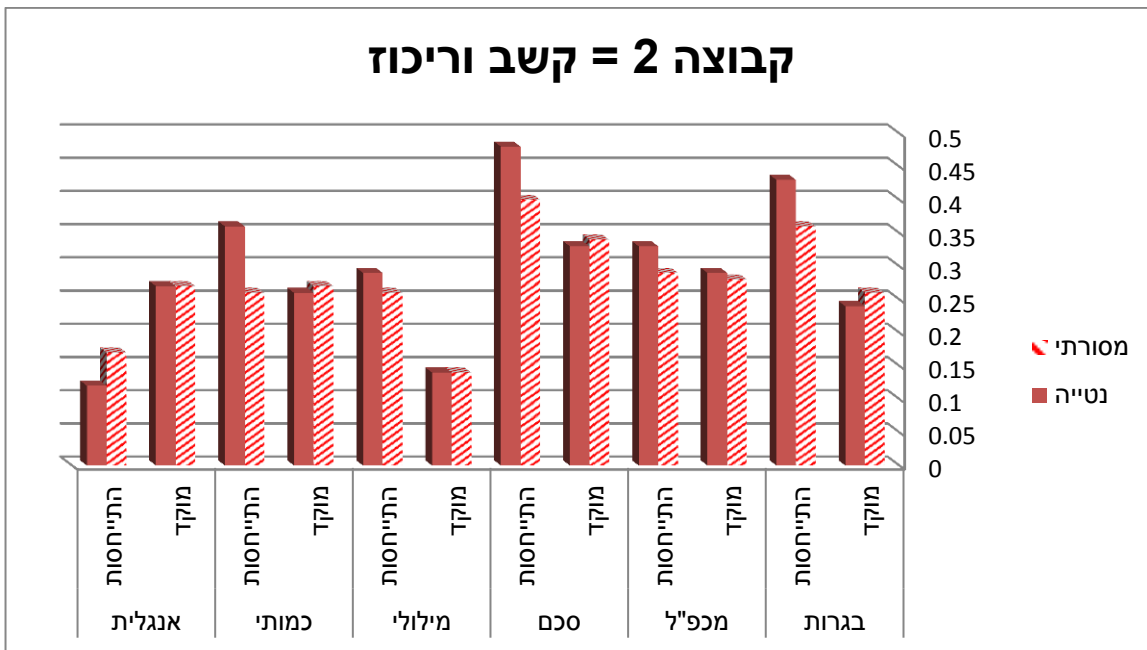
מהתבוננות באיור 6 עד איור 10 ניתן לראות כי (פרט לאנגלית קבוצה 98 ובגרות קבוצה 99) גודל האפקט במחקר ציון הנטייה קטן בערכו המוחלט (ובדרך כלל שווה סימן) מזה שבמחקר המסורתי.

5.3.3. תוקף דיפרנציאלי

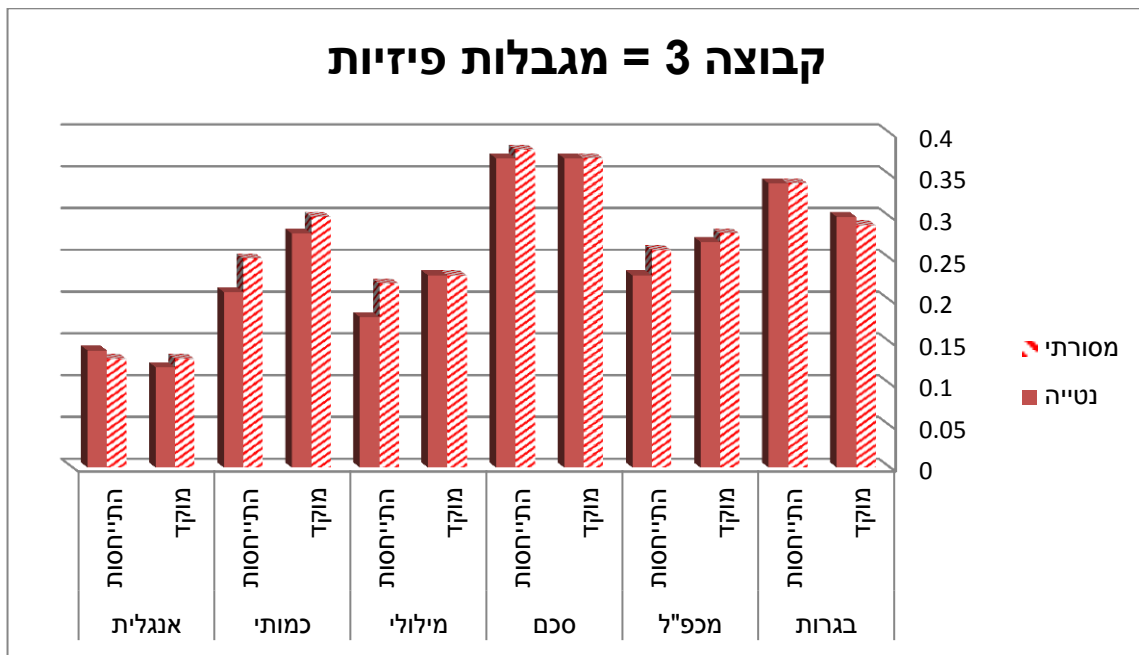
איור 11 - מקדם תוקף עבור קבוצה 1 = לקויות למידה



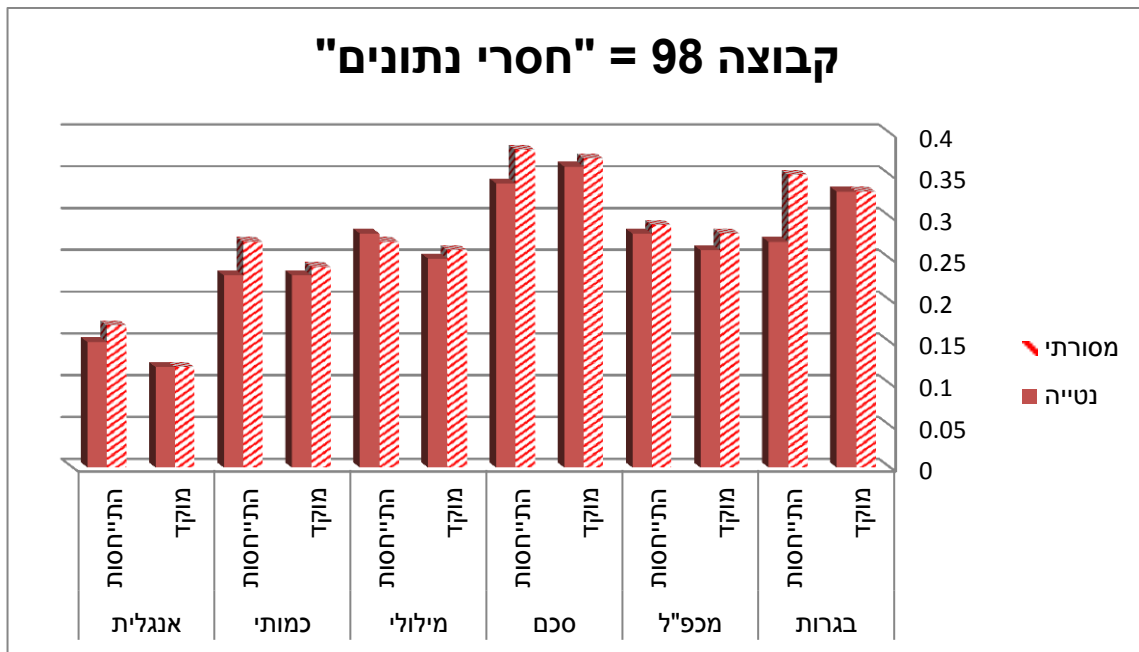
איור 12 - מקדם תוקף עבור קבוצה 2 = קשב וריכוז

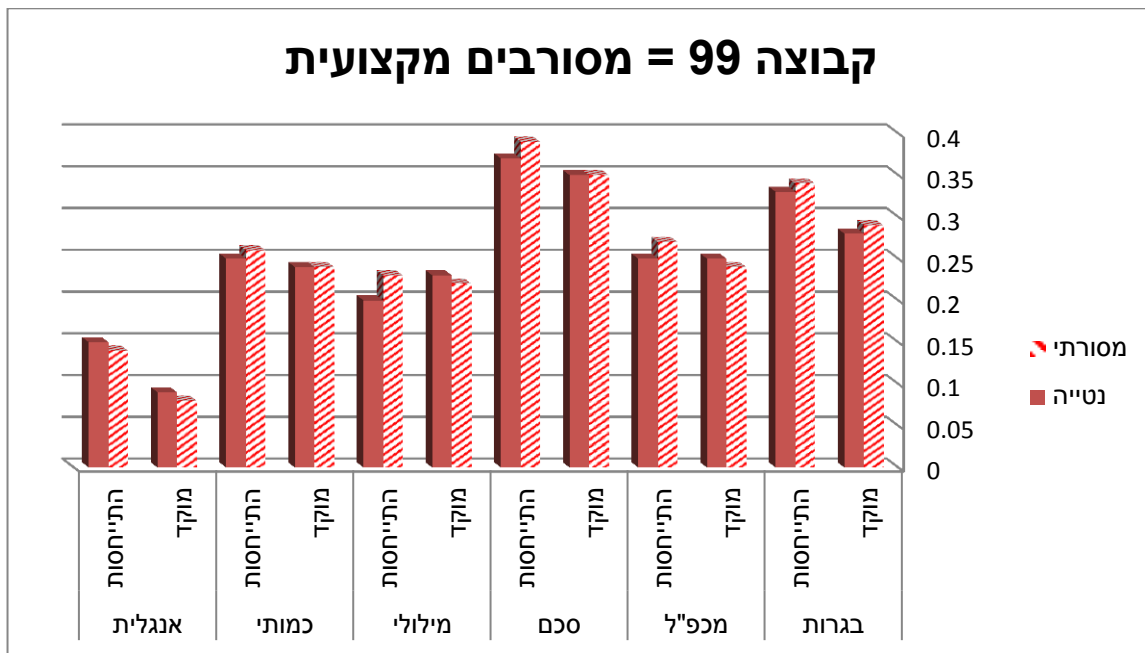


איור 13 - מקדם תוקף עבור קבוצה 3 = מגבלות פיזיות



איור 14 - מקדם תוקף עבור קבוצה 98 = "חסרי נתונים"





מהתבוננות באיור 11 עד איור 15 ניתן לומר ששתי שיטות המחקר, המחקר המסורתי והמחקר בשיטת ציון הנטייה, נותנות כמעט אותו מקדם תוקף. ראוי לציין שלגבי קבוצות המוקד אנו מצפים שהתוצאות תהיינה דומות; זאת משום שקבוצת מוקד בהשוואה מסורתית ואותה הקבוצה בהשוואה בשיטת ציון הנטייה למעשה אמורות להכיל כמעט אותם האנשים. (בהקשר זה נשים לב לכך שכאשר התוקף מחושב, הוא מחושב על ידי חישוב מקדם מתאם בין החזאי לקריטריון עבור כל אחת מקבוצות המוקד וההתייחסות בנפרד). ההבדל הקטן שכן נוצר הוא מסילוק אנשים עם נתונים חסרים. לגבי קבוצת ההתייחסות, שעבורה התוקף במחקר בשיטת ציון הנטייה מחושב עבור קבוצה קטנה מאוד יחסית לקבוצה המקורית (ובהתאם לכך גם יחסית לקבוצה בשיטת המסורתית), העובדה שהתוקף יוצא די דומה בשתי שיטות המחקר מראה שהבחירה שנעשתה על ידי ציון הנטייה אכן מייצגת את האוכלוסייה כולה בצורה טובה.

- אנגל, א' (2007). השפעת תוספת זמן בבחינה הפסיכומטרית על נבחנים עם הפרות קשב וריכוז. *עבודת גמר לתואר מ.א.*, ביה"ס לחינוך, האוניברסיטה העברית בירושלים.
- מושינסקי, א' (2000). *סיכום תוצאות ניסוי מפע"מ* (דוח מס' 276). ירושלים: מרכז ארצי לבחינות ולהערכה.
- נבון, ד' וכהן, י' (2009). אפליה בבחינות הכניסה לאוניברסיטאות או ארטיפקטים סטטיסטיים: תגובה לגמליאל וקאהן (2004). *מגמות, כרך מו'*, עמ' 398-418.
- 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.
- קנת-כהן, ת', ברונר, ש' ואורן, כ' (1999). ניתוח-על של תוקף הניבוי של מרכיבי מערכת המיון לאוניברסיטאות בישראל כלפי מידת ההצלחה בלימודים. *מגמות, כרך מ' מספר 1*, עמ' 54-71.
- קנת-כהן, ת., טורוול, א. ואורן, כ. (2011). *הוגנות מערכת המיון לאוניברסיטאות לפי שפת היבחנות במכפ"ל (ערבית או עברית) ומגדר* (דוח מס' 376). ירושלים: מרכז ארצי לבחינות ולהערכה.
- American Educational Research Association, American Psychological Association & National Council on Measurement in Education (1999). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.
- Bennet, R.E., Rock, D.A., & Jirele, T. (1986). *The Psychometric Characteristics of the GRE General Test for Three Handicapped Groups* (ETS research Report RR- 86-6). Princeton, NJ: Educational Testing Service.
- Braun, H., Ragosta, M. & Kaplan B. (1986). *The Predictive Validity of the Scholastic Aptitude Test for Disabled Students* (Research Report 86-38). New York: College Entrance Examination Board.
- Camara, W.F., Copeland, T. & Rothchild, B. (1998). *Effects of Extended Time on the SAT I: Reasoning Test: Score Growth for Students with Learning Disabilities* (College Board Research Report No. 1998-7). New York: College Board.
- Camilli, J. (2006). Test Fairness. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 221-256). Westport: American Council on Education & Praeger.
- Centra, J.A. (1986). Handicapped Student Performance on the Scholastic Aptitude Test. *Journal of Learning Disabilities*, 19, 324-327.
- 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* (pp. 24-25). New York: Lawrence Erlbaum associates.
- Cohen, Y., Ben-Simon, A., Moshinsky, A., Eitan, M. (2008). *Computer Based Testing (CBT) in the Service of Test Accommodations*. (Report No. 348). Jerusalem: National Institute for Testing and Evaluation.
- D'agostino, R.B. (1998). Tutorial in Biostatistics Propensity Score Methods for Bias Reduction in the Comparison of a Treatment to a Non-Randomized Control Group. *Statist. Med* 17, 2265-2281.
- Darlington, R.B. (1971). Another Look at "Cultural Fairness". *Journal of Educational Measurement*, 8, 71-81.
- Diagnostic and Statistical Manual of Mental Disorders DSM-IV-TR Fourth Edition (2000). *American Psychiatric Association; 4th edition*.
- Flanagan, D.P., Fiorello, C.A., & Ortiz, S.O. (2010). Enhancing Practice through Application of Cattell-Horn-Carroll Theory and Research: A "Third Method" approach to specific learning disability identification. *Psychology in the Schools*, 47(7), 739-760.
- Fletcher, J.M., Foorman, B.R., Boudousquie, A., Schatschneider, C., & Francis, D.J. (2002). Assessment of Reading and Learning Disabilities: A research-based intervention-oriented approach. *Journal of School Psychology*, 40, 27-63.
- Gershman, F.M., MacMillan, D.L., & Bocian, K.M. (1996). Learning Disabilities, Low Achievement, and Mild Mental Retardation: More alike than different? *Journal of Learning Disabilities*, 29, 570-581.
- Laing, J. & Farmer, M. (1984). *Use of the ACT Assessment by Examinees with Disabilities* (Research Report No. 84). Iowa City, IA: American College Testing.
- Linn, R.L. (1983a). Pearson Selection Formulas: Implications for Studies of Predictive Bias and Estimates of Educational Effects in Selected Samples. *Journal of Educational Measurement*, 20, 1-16.
- Linn, R.L. (1983b). *Predictive Bias as an Artifact of Selection Procedures*. In H. Wainer & S. Messick (Ed.), *Principals of modern psychological measurement: A Festschrift for Frederic M. Lord* (pp. 27-40). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Linn, R.L., & Hastings, C.E. (1984). Group Differentiated Prediction. *Applied Psychological Measurement*, 8, 165-172.
- Linn, R.L., & Werts, C.E. (1971). Considerations in Studies of Test Bias. *Journal of Educational Measurement*, 8, 1-4.
- Linn, R.L. (1984). Selection Bias: Multiple Meanings. *Journal of Educational Measurement*, 21, 33-47.

- Lord, F.M. & Novick, M.R. (1968). *Statistical Theories of Mental Test Scores*. Addison-Wesley Publishing Company, Inc.
- Lovett, B.J. (2010). Extended Time Testing Accommodations for Students with Disabilities: Answers to five fundamental questions. *Review of Educational Research*, 80, 611-638.
- MacArthur, C.A. (1996). Using Technology to Enhance the Writing processes of Students with Learning Disabilities. *Journal of Learning Disabilities*, 29, 344-354.
- Moshinsky, A., Kazin, C. (2004). *Constructing a Computerized Psychometric Adaptive Test for University Applicants with Disabilities*. (Report No. 315). Jerusalem: National Institute for Testing and Evaluation.
- Oren, C., & Even, A. (2005). *The Fairness and Validity of the Higher Education Selection System for Students with Disabilities* (Report No. 325). Jerusalem: National Institute for Testing and Evaluation.
- Phillips, S.E. (1994). High Stakes Testing Accommodations: Validity versus Disabled Rights. *Applied Measurement in Education*, 7, 93-120.
- Pitoniak, M.J., & Royer, J.M. (2001). Testing Accommodations for Examinees with Disabilities: A review of psychometric, legal and social policy issues. *Review of Educational Research*, 71(1), 53-104.
- Rosenbaum, P. R., & Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, 70, 41-55.
- Scanlon, D. (2013). Specific Learning Disability and Its Newest Definition: Which Is Comprehensive? and Which Is Insufficient? *Journal of Learning Disabilities*, 46(1), 26-33.
- 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.
- Tindal, G., & Fuchs, L. (1999). *A summary of Research of test Accommodations: What We Know So Far*. Mid-South Regional Resource Center, University of Kentucky.
- Wightman, L.F. (1993). *Test Takers with Disabilities: A Summary of Data from Special Administrations of the LSAT* (Research Report 93-03). Newtown, PA: Law Admission Council.
- Willingham, W.W., Ragosta, M., Elliot Bennett, R., Braun, H., Rock, D.A., & Powers, D.E. (1998). *Testing Handicapped People*. Allyn and Bacon, MA.
- 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.

Ziomek, R.L. & Andrews, K.M. (1996). *Predicting the College Grade Point Averages of Special-Tested Students from their ACT Assessment Scores and High School Grades* (ACT Research Report Series, 96-7). Iowa City, IA: American College Testing.

Zurcher, R., & Bryant, D.P. (2001). The Validity and Comparability of Entrance Examination Scores after Accommodations are made for Students with LD. *Journal of Learning Disabilities*, 34(5), 462-471.

7.1. מודלים ושיטות

7.1.1. ה- Δ של כהן

המדד Cohen's d לגודל אפקט הוא הפער המתוקן בין ממוצעי שתי קבוצות במשתנה רציף. d מוגדר כממוצע קבוצת המוקד (f) פחות ממוצע קבוצת ההתייחסות (r) חלקי ממוצע סטיות התקן במשתנה הנדון,

$$d = \frac{\bar{X}_f - \bar{X}_r}{\sigma} \text{ כאשר } \sigma = \sqrt{\frac{\sum_{i=1}^{n_f} (x_{f,i} - \bar{x}_f)^2 + \sum_{i=1}^{n_r} (x_{r,i} - \bar{x}_r)^2}{n_f + n_r}}$$

קיימת נוסחה שקולה מתמטית להגדרה הקודמת בה מחושב 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 הנוסחה²²:

$$d = r_{gv} / \sqrt{pq(1-r_{gv}^2)}$$

כאשר r_{gv} הוא מקדם המתאם של פירסון בין משתנה השיוך הקבוצתי (g) המוגדר כ-0 עבור קבוצת ההתייחסות וכ-1 עבור קבוצת המוקד לבין המשתנה הנדון (v) והמשתנים p ו- q הם הפרופורציות של כל אחת משתי הקבוצות בצירוף של שתיהן. הנוסחה השנייה שימושית בעיקר לצורך חישוב הערך המתוקן לקיצוץ תחום/למהימנות של d , כאשר אז מציבים בה את ה- r המתוקן לקיצוץ התחום/למהימנות²³.

הפרש d -ים של כהן (Thorndike, 1971) המסומן ב- Δ הוא ה- d בחזאי פחות ה- d בקריטריון $\Delta = d_p - d_c$. ערך זה נקרא על ידנו "ה- Δ של כהן". ערך חיובי של הפרש זה מבטא הטיה לטובת קבוצת המוקד וערך שלילי הטיה לרעתה.

7.1.2. המודל של קלירי

משוואת רגרסיה אחת (מעבר לשתי הקבוצות) מחושבת עבור חיזוי הקריטריון (C) על ידי החזאי (P) בהתאם למשוואות הבאות:

$$\hat{C} = aP + B \text{ כאשר } a = \frac{Cov(P,C)}{Var(P)} = r_{PC} \cdot \frac{s_C}{s_P} \text{ , } b = \bar{C} - a\bar{P}$$

עבור כל תצפית, ערך הקריטריון בפועל הוחסר מערך הקריטריון החזוי לפי קו הרגרסיה המשותף לפי הנוסחה: $\hat{C} - C =$ שארית. השארית ממוצעת לקבלת מדד קלירי.

²² ראה פיתוח של הנוסחה, בהתבסס על כהן (Cohen, 1988), אצל קנת-כהן (קנת-כהן, 2001).
²³ ראה הערת שוליים מספר 19.

7.1.3 . מקדם תוקף

מקדם התוקף הוא מקדם המתאם של פירסון בין החזאי (P) לקריטריון (C) ומחושב (עבור קבוצת המוקד וההתייחסות בנפרד) בהתאם לנוסחה הבאה:

$$r_{PC} = \frac{Cov(P, C)}{\sigma_P \cdot \sigma_C} = \frac{\sum_{i=1}^n (P_i - \bar{P}) \cdot (C_i - \bar{C})}{\sqrt{\sum_{i=1}^n (P_i - \bar{P})^2} \cdot \sqrt{\sum_{i=1}^n (C_i - \bar{C})^2}}$$

7.1.4 . ציוני נטייה

בניסוי במערך אקראי, האקראיות בהקצאת האנשים לקבוצות טיפול שונות מבטיחה שבממוצע לא יהיה הבדל שיטתי במשתני הרקע (covariates) (בין אם הם נמדדים או סמויים). אבל, במחקר תצפיתי (לא במערך אקראי), לחוקרים אין שליטה על שיבוץ הטיפולים ולכן השוואה ישירה של התוצאות בין קבוצות הטיפול השונות יכולה להיות מטעה. קושי זה יכול להימנע, לפחות באופן חלקי, אם מידע על משתני הרקע הנמדדים משולב לתכנון המחקר. שיטות מסורתיות של תיקון בדרך כלל מוגבלות מכיוון שהן יכולות לכלול מספר מוגבל בלבד של משתני רקע לצורך התיקון. אבל, ציון נטייה (propensity score), אשר מספק מספר סקלארי אחד המסכם את כל המידע הנמצא במשתני הרקע, ובכך למעשה מדמה הקצאה מקרית, איננו בעל המגבלה הזאת.

באופן אינטואיטיבי, ציון הנטייה הוא מידת הסבירות שאדם יקבל טיפול (ישתייך לקבוצת המוקד) בהינתן רק משתני הרקע שלו (D'agostino, 1998). באופן פורמלי (Rosenbaum & Rubin, 1983), ציון נטייה למשתתף i מוגדר כהסתברות המותנית לקבלת טיפול מסוים ($T_i = 1$) לעומת אי-קבלת הטיפול ($T_i = 0$) בהינתן וקטור \vec{x}_i המהווה את וקטור משתני הרקע שלו

$$propen(\vec{x}_i) = Prob(T_i = 1 | \vec{X}_i = \vec{x}_i)$$

כאשר מניחים שבהינתן ה- X ים, ה- T ים הם בלתי-תלויים

$$Prob(T_1 = t_1, \dots, T_N = t_N | \vec{X}_1 = \vec{x}_1, \dots, \vec{X}_N = \vec{x}_N) = \prod_{i=1}^N propen(\vec{x}_i)^{t_i} \cdot (1 - propen(\vec{x}_i))^{1-t_i}$$

מהגדרה זו נובע כי T ו- X בלתי תלויים בהינתן $propen(x)$; ולכן, אנשים בקבוצת המוקד וההתייחסות בעלי אותו ציון (או ציון קרוב) יטו להיות בעלי התפלגות דומה במשתני הרקע.

במידה ומשתני הרקע לא מכילים מידע חסר, ניתן להעריך את ציון הנטייה על ידי רגרסיה לוגיסטית.

7.2. הרחבה על שתי השורות הראשונות בלוח 18

בתת-פרק זה ננסה להסביר את הממצאים שהתקבלו עבור מדד ה- Δ של כהן עבור שתי הקבוצות בעלי לקויות-למידה ובעלי בעיות קשב וריכוז.

ביתר פירוט: אנו מנסים להסביר מדוע התקבלו תוצאות שונות עבור המחקר המסורתי ועבור המחקר בעזרת ציון הנטייה הן בבגרות והן במכפ"ל, ומדוע כאשר משווים את התופעה עבור הבגרות ומכפ"ל מתקבלת תופעה שהיא לכאורה הפוכה.

הממצאים הרלוונטיים מסוכמים בלוח הבא (בהתאם לכלל האצבע של כהן, 1998), ומהווים ממצאים חלקיים לתוצאות המלאות המופיעות בדו"ח²⁴.

לוח 21 - ה- Δ של כהן עבור חזאים מרכזיים ושתי הקבוצות הראשונות

קבוצה						חזאי					
						מכפ"ל		בגרות		סכם	
						DP	DT	DP	DT	DP	DT
1 = לקויות למידה						-	-	-	-	-	-
2 = קשב וריכוז						-	-	-	-	-	-

? מדוע יש ערך שלילי (-) בכהן מסורתי עבור החזאי בגרות?

קבוצת האנשים שמקבלים תנאים מיוחדים נקבעת לפי מכפ"ל. מאוד יתכן שקבוצה זו איננה זהה לקבוצת מקבלי התנאים בבגרות, אלא מהווה קבוצה חלקית לה. לכן, בבגרות ישנם אנשים שמקבלים תנאים אבל נמצאים בקבוצת ה"רגילים" אליהם אנו משווים; מה שאומר שקבוצת ההתייחסות אליה אנו משווים בבגרות חזקה יותר ולכן נוצר ניבוי חסר בבגרות.

✓ טיעון זה איננו תקף כאשר משתמשים בשיטת ציון הנטייה בבגרות היות ואז קבוצת ההתייחסות נבחרת בהתאם להשוואת ציוני הנטייה שמתחשב בגיל ומין, כך שהאפקט של יותר אנשים חזקים מתבטל.

✓ כמובן שהטיעון לא תקף גם עבור מכפ"ל בשיטת כהן מסורתי שעבורו בוחרים את קבוצת מקבלי התנאים בצורה המתאימה להשוואה.

? מדוע אם כן יש ערך שלילי (-) בכהן שיטת ציון נטייה עבור החזאי מכפ"ל?

²⁴ ראה לוח 20 למקרא של הסימונים.

כאשר משתמשים בציון הנטייה על מנת לבחור את קבוצת ההתייחסות אנו עושים זאת בהתאם לאנשים בקבוצות הלקות שעבורן גיל הנבחנים מבוגר יותר (+1.06/+1.42 שנים בהתאמה יחסית לקבוצת ה"רגילים") ופרופורציית הגברים גבוהה יותר (+7%/+19% בהתאמה יחסית לקבוצת ה"רגילים"). (ראה בלוח הבא).

לוח 22 - התפלגות מין וגיל לפי קבוצות (0,1-2)

גיל			מין			קבוצה
ממוצע	אנשים	סטטיית תקן	אחוז גברים	אחוז נשים	אנשים	
21.51	119,526	2.49	42%	58%	111,037	0 = רגילים
22.57	940	2.07	49%	51%	927	1 = לקויות למידה
22.93	187	2.29	61%	39%	182	2 = קשב וריכוז

מהסתכלות בשני הלוחות הבאים (אשר הנתונים בהם מחושבים עבור אוכלוסייה הזוהה לאוכלוסיית המחקר שלנו ומחושבים מעבר לקבוצות מקבלי התנאים), ניתן לראות שנבחנים מבוגרים יותר מצליחים פחות באוניברסיטה וכן שגברים מצליחים פחות באוניברסיטה. כך, שעל פניו, יש לנו כאן סתירה שכן אנו מחליפים את קבוצת ההתייחסות בקבוצה יותר חלשה בשני המקרים.

לוח 23 - ממוצע וסטטיית תקן של הקריטריון ושני החזאים המרכזיים לפי גיל

בגרות		מכפ"ל		ציון שנה א'			קבוצת גיל (שנים)
ממוצע	סטטיית תקן	ממוצע	סטטיית תקן	ממוצע	סטטיית תקן	אנשים	
99.01	7.75	635.72	81.62	82.88	9.04	6,245	21.25-21.50
98.85	7.73	638.98	79.95	82.38	9.17	6,533	21.50-21.75
98.56	7.83	640.49	79.12	82.45	9.20	6,260	21.75-22.00
98.01	8.08	638.53	81.64	81.9	9.38	5,625	22.00-22.25
97.40	7.95	637.84	82.13	82.00	9.46	5,078	22.25-22.50
97.23	7.96	641.74	78.68	81.73	9.38	4,994	22.50-22.75
97.04	8.17	638.20	83.92	81.58	9.72	4,470	22.75-23.00

לוח 24 - ממוצע וסטטיית תקן של הקריטריון ושני החזאים המרכזיים לפי מין

בגרות		מכפ"ל		ציון שנה א'			מין
ממוצע	סטטיית תקן	ממוצע	סטטיית תקן	ממוצע	סטטיית תקן	אנשים	
98.51	9.14	653.88	76.00	80.88	10.21	48,090	זכר
99.29	8.53	612.87	86.96	83.14	8.99	66,117	נקבה

כדי לקבל הסבר לתופעה יש לבחון את הנתונים יותר לעומק.
ניתן לראות בלוח 23 שפחות או יותר אין שינוי במכפ"ל עם הגיל, לעומת זאת, ציון שנה א' נמצא במגמת ירידה עם עליית הגיל. נתון זה אומר שכאשר אנו מחליפים את קבוצת ההתייחסות בקבוצה מבוגרת יותר אנו מחליפים בקבוצה שעבורם ציון מכפ"ל יטה לתת הערכת יתר. אפקט דומה קורה עבור מין: בעוד שהמעבר מנקבה לזכר גורם לציון מכפ"ל לגדול, ציון שנה א' לעומתו דווקא קטן, מה ששוב אומר שמעבר זה כרוך במעבר לקבוצה עם הערכת יתר. כאשר מחליפים את קבוצת ההתייחסות בקבוצה אחרת שיחסית אליה נמצאת בהערכת יתר סביר הוא שעבור הקבוצה המשווית תתקבל כעת הערכה נמוכה יותר שבמקרה שלנו הופכת להערכת חסר.

לבסוף, נעיר שטיעונים דומים לעיל (רק טיפה יותר מורכבים שכן יש לחשב את השינוי במונחי סטיות התקן ולהתחשב במספר האנשים בקבוצות) יתמכו שוב בהבדל הקיים במדד ה- Δ של כהן בין השיטה המסורתית לזו של שיטת ציון הנטייה עבור הבגרות (וזאת בנוסף לטיעון הקודם בתחילת תת-פרק זה).