

לוח המרכז

335

אפליה או
ארטיפיקטים סטטיסטיים?
תגובה לגמליאל וקאהן (2004)

דוד נבוין
יעקב כהן

דצמבר 2006



מרכז ארכי לבחינות ולהערכת (ע"ר)
NATIONAL INSTITUTE FOR TESTING & EVALUATION
מייסודה של האוניברסיטה הפתוחה בישראל
ת"ד 26015 ירושלים 91260 טל' 02-6759555

דוח מרכז 335

ISBN:965-502-132-7

© All rights reserved
NITE
P.O.B. 26015 Jerusalem

© כל הזכויות שמורות
מרכז ארכאי לבחינות ולהערכה
ת"ד 26015 ירושלים 91260

**אפליה או ארטיפקטים סטטיסטיים?
תגובה לגלילי וקאהן (2004)**

דוד נבון יואב כהן

תקציר

גמליאל וקאהן (2004) טוענים שהוגנות הילכית ביריה מותנה בכך שיישמר בקרבת המתකבים הייצוג היחסני שהיה לקבוצות מיעצב סוציאו-אקונומי אילו נשתה הקבלה על-פי חזאי בעל תקפות מושלמת. עוד הם טוענים שתנאי זה אינו מתאפשר מתמטית עם התנאי המקובל להוגנות הנקרה "חוסר-הטיה בניבוי", ועל כן ממצאים המצביעים על ההתקומות בפערים בין הקבוצות מהחזאי לקריטריון (כלומר, שהפער המתוקן בין הקבוצות במומצאי הציון המשמש ליריה לאוניברסיטאות בארץ גבוה מהפער המקורי במומצאי הקריטריון) מהווים ראייה להטיה תרבותית-חברתית של הליך היריה כנגד קבוצת המיעצב הסוציאו-אקונומי הנמור. כנגד טיעונים אלה מובאים במאמר זה מספר טיעוני נגד. העיקריים בהם הם : א. תנאי ההוגנות המוגדר ע"י ייצוג יחסית הינו חמור מכדי להיות ישים, ב. הטענה לאי-התישבותו עם תנאי חוסר-הטיה בניבוי מבוססת על הגדרה מצטמת מודד של האחרון, ג. פערים בין מומצאי קבוצות תלויים מדי במספר גורמים סטטיסטיים מכדי לשמש כראיה להטיה מובנה, ד. כאשר מביאים בחשבון גורמים אלה בניתוח ממצאים המצביעים על ההתקומות בפערים המתוקנים בין הקבוצות מהחזאי לקריטריון, נראה שהם מסבירים את ההתקומות.

גמליאל וקהאן (2004) טוענים שציוני הקבלה לאוניברסיטאות מפלים לרעה קבוצות חלשות מבחינה מעמדן הסוציאו-כלכלי. הבעה במסקנה זו, שהטיעונים המשמשים את המחברים לביסוסה לוקים במספר חולשות.

ההגדרה

הוגנות חברתית של הילך ביריה מוגדרת על-פי התאמת שיעורי הקבלה נצפים של קבוצות אוכלוסייה שונות לשיעורים הנקבעים על פי מודל אידיאלי מסוים (למשל, שיעורים שוים). בספרות הוציאו הגדרות Cole & Zieky, 2001; Darlington, 1971; Hunter, Schmidt, & Rauschenberger, 1977; Linn, 1984; Peterson & Novick, 1976; Willingham & Cole, 1997 (Cahan & Gamliel, 2001) אחרות הם מכנים "התיה בברירה הכרחית". הטיה כזו היא מתקימת לפי הגדרת המחברים בכל מצב בו שיעור המתקבלים מקבוצת אוכלוסייה נתונה אינו שווה לשיעור שהוא מתתקבל מאותה קבוצה אילו באורך-פלא היה ניתן לדעת במדוק מה תהיה הצלחתו בלימודים של כל מועמד, אף כזו שבפועל לא התקבל.

דא עקא, אי-שוויון כזו מתקיים כמעט בכל הילך ביריה. כאשר בוררים על-פי חזאי הצלחה שתקופתו אינה מושלמת, גודל ההפרש בין השיעורים תלוי במידה בה הפער בין ממוצעי שתי הקבוצות על המדר המשמש כחזאי שונה בין ממוצעיהם על מדר הצלחה בלימודים. נדר ביותר המקורה בו הפערים שוים לחלוtin. הסיבה לכך פשוטה: לאחר שמדד הטעב המתאים בין מדר קלשאנו המשמש למינן לבין הצלחה בלימודים איןנו מלא, לא סביר שהמתאים בין כל אחד מהם לבין משתנה שלishi יהיו שוים. מכיוון שכאשר המשתנה השלישי הוא אפיון דיקוטומי כלשהו של האוכלוסייה, לא סביר שהיה שוויון מלא בין הפער בממוצעי מדר המינן יהיה גדול יותר וייתכן שייהיה קטן יותר. במקרה הראשון הקבוצה תמחה מז בממוצעי מדר המינן יהיה גדול יותר וייתכן שייהיה קטן יותר. במקרה השני הקבוצה תמחה מז הסטם על אףיה, אך במקרה השני יהיה לכאורה בסיס לקבוצה החזקה למחרות על אףיה. אלא שבשני המקרים אין בהכרח אףיה. בסיס לטענת אףיה קיים כשהטהה נובעת משימוש במדד מינן שמוספעים מוגרמים תרבותיים-חברתיים בלתי-תלויים בכשור אקדמי ([להלן הטיה מובנית](#)). למרות שקיימות אפשרויות כזו, הבדלי פערים הם תופעה שכיחה שצומחת פשוט מהמציאות המוכרת בברירת מועדים: חייזוי הצלחה איןנו מדויק ועל-כן המינן אינו מושלם. גם אילו היה במקרה שוויון בפערים בין קבוצות אוכלוסייה א' ו-ב', (למשל, דובר עברית כשפת א' ודובר עברית כשפת ב'), סביר מדוע שני שטי קבוצות אוכלוסייה אחרות, ג' ו-ד' (למשל, גברים ונשים), שלא היה מתקיים שוויון בפערים ביניהם. מארח שמספר החלוקת השונות של אוכלוסייה הוא רב, מעשית אין בנמצא ביריה בה לא תתקיים כלל הטיה כלשהי לפי הגדרה מרוחיבה זו.

האם "התיה" כזו אינה הוגנת? כמובן, הוגנות אינה מסתמכה בייצוג הולם לקבוצות אוכלוסייה. גם מעבר לקושי המזוכר לעיל בהשגת ייצוג כזה בשל קשרים אקולוגיים בין משתני קבוצה שונים (כמו מגדר, מיצב סוציאו-כלכלי, ארץ לידה, וכו'')³, קיימים קושי עקרוני בנסיבות מסווג ההוגנות לאי-התיה כנדג' קבוצה חברתית מסוימת. יסודות המשפט המודרני מכונים בעיקר להבטחת זכויותיו של הפרט. מקובל אפילו להימנע ככל האפשר מפגיעה בזכויותיהם של פרטיהם ללא קשר להשתיכותם הקבוצתית. על-כן, במידה שסביר להציג את המוקומות המוגבלים במסלול לימודים נתון על-פי סיכון הצלחתו של כל מועמד ומועד (גישה המכונה "מריטוקרטית"), ספק רב אם הוגן לדחות מועמד שסיכון הצלחתו לטובים לטובה מועמד שסיכון פחותים בשל שיקולי השתיכות קבוצתית. כמה פסיקות שניתנו בארה"ב בעשור האחרון

הצביעו על הביעתיות המשפטית בהלכתי קבלה המותחים עד כדי כך את הניסיוןקדם קבוצות אוכלוסייה חלשות (ראה סקירה ב- Alba et al, 2002).

די בשיקולים עקרוניים אלה לא רק כדי לבדוק את נקודת המוצא של הגישה המריטופרטית, אלא גם כדי לפקפק בסבירותו של כל מילון המעניק קידימות לייצוג הולם על-פני תוקף ניבוי. אולם, כאשר נחיזית אפשרות להטיה כנגד קבוצה שיש אינטרס חרמתי להגן עליה, מתקבל לתהויחת לתופעה (בעיקר בתחום של קבלה לעובדה)قبال השפעה מזיקה (adverse impact) שיש לננות לצמצמה באמצעות העדפה מותקנת. אך

הסיג המقبول הוא שניתן לעשות זאת מבלי לפגוע בתקפות (Equal Employment Opportunity), למשל לבровер על-פי השתייכות קבוצתית כאשר נתוני הקבלה שוים. הנהל הקיצוני יותר, של מתן קידימות לשתייכות קבוצתית על-פני סיכון הצלחה, גולש כבר לאפליה, קרי קיפוח פרטיהם בשל השתייכותם הקבוצתית. כאמור, ספק אם מולך כזה ראוי, אף כשהוא נעשה למען מטרת הנטפסת קרואה.

במילון להשכלה גבוהה נהוג לחוש לחטיה ממשית רק כאשר אי-השוויון בפערים עשוי להיגרם מכש של מדדי המילון נטיה שיטית לאומדן-חסר של ה联系ים האקדמיים של חברי הקבוצה החלשה. חשש זה אינו מוצק כל עוד הפרשים בפרמטרים של משווהות הניבו לשתי הקבוצות אינם חריגים מתנאי גבול מסוימים (Birnbaum, 1979, 1981; Linn, 1984).

מאחר שהצעתם של המחברים קבוע מקרים קבלה דיפרנציאליות לחבריו קבוצות שונות (גמליאל וכהאן, ע' 443) אינה מתחשבה בסיג זה, סביר בויתר שתתגע בתקפות. כל מועמד מקבוצה א' המתבל בשאל מבחן קבלה גבוהה ייחסת על-חשבון מועמד מקבוצה ב' שצינו טוב יותר במידד תקף המשמש לקבלה, תורם מן הסתם להפחחת התקפות. על-כן, ניתן לדחות את הצעת המחברים¹ על פניה כשם שנדחו הצעות דומות בעבר (ראה, למשל, הדיוון ב-Darlington, 1971).

עם זאת, מטרתו העיקרית של מאמנו לדון בטיעוניהם של המחברים לגופם. ספציפית, עד כמה טיעוניהם קבילים מנוקדת ראותו של מי שבשל שיקולי הערכיהם איננו מאמין בקידומתו של העיקרונו המריטופרטטי?

כזכור, הגדרתם של המחברים להונאות הלכתי ביריה דורשת ששיעוריו המתקבלים בקרב קבוצות מיצב סוציאו-כלכלי שנותו ישתוו אלה שהיו נצפים אילו נעשתה הקבלה על-פי חזאי בעל תקפות מושלמת. היא רואה אפוא הטיה בקיומו של פער בין הקבוצות במומחיות הצלחה בלימודים הקטן במידה משמעותית מהפער ביןין במומחיות המדד המשמש לקבלה. דא עקא, לא-שוויון כזה בפערים עשויות להיות סיבות פסיכומטריות תמיינות למדוי.

למשל, מאחר שב"כ למדד הצלחה בלימודים מהימנות נמוכה מזו של המדדים המשמשים לחיזוה, צפוי שפער בין קבוצות כלשהן עליו יטה להיות קטן מפערים ביןין על אחד מהחזאים, בשל התופעה המוכרת של הנחתת מתאימים (ראה למשל, Guilford, 1965/1942; Linn, 1973). לדוגמה, אם מהימנות הקרייטוריון של הצלחה בלימודים באוניברסיטה (כפי שהיא נאמדת ע"י המתאם בין שני חצאים

1. הצעה אחרת שלהם – לקבל על-פי חזאים המבוססים על הערכות מורים בלימודים קודמים – לוקה בחיסרונו דומה, שכן תקפותם של חזאים אלה נמוכה בעיקר משור הסטנדרטיות שלהם. יתר על כן, הניסיון בקבלת אוטומטיות של מצטיינים בבתי ספר תיכוניים שהציגו באה"ב מלמד שיטתה זו אף אינה מועילה כמעט להגברת השוויון הבין-קבוצתי במסדות אקדמיים (ראה סקירה ב-Alba et al, 2002).

אקראים של ציוני שנה א') היה כ- 0.68 ב ממוצע (ראה קנט-כהן, ברונר ואורו, 1999, ע' 61), הרי שהיא פחותה ב- 29% לערך ממהימנות הבדיקה הפסיכומטרית בשלושת המועדים האחרונים (כ- 0.96; סטולר, 2005). במלים אחרות, משקל גורם הטעות בשונות הקריטריון רב יותר מאשר משקלו בשונות החזאי. תיקון שונות הקריטריון עבר הנחתה זו מקטין את המכנה בחישוב הפערים בין הקבוצות על פני הקריטריון, כך שלאחר התיקון מנת הפערים קטנה ב- 16% לערך מהמנה שחושבה לפני התיקון. הדגמה זו לוגaldo האפשרי של אפקט המהימניות (או של כל גורם פסיכומטרי אחר) אינפורמטיבית בפניהם עצמה. אך היא אולי חשובה יותר כתזכורת לעצם העובדה שסיכון הגדרת המחברים להוגנות תלוי בקיומו של אפקט פסיכומטרי זהה או אחרים. נניח מצב היפוטטי בו גם מהימנות החזאי הייתה 0.65 משותחת של הבדיקה הפסיכומטרית הייתה קפדנית פחות מאשר היא היום. ברור שהיה בכך כדי להקטין את אי-השוון בפערים (ועל-כן גם בשיעורי הקבלה המופיעים בהגדרת ה"טיה בברירה הכרחית"). אך האם סביר לומר שבכך היה המבחן הופך להוגן יותר? דוגמה זו ממחישה אפוא את הביעתיות הבסיסית שהגדירה: אין היגיון רב באבחון הטיה הנובעת מחוסר הוגנות על-פי שוון בין גדלים המושפעים מן הסתמיים משתנים פסיכומטריים ואחרים שקשה לדעת את ציוני השפעותיהם. להלן, בפרק "פירוש הממצאים האמפיריים", נדונים עוד גורמים פסיכומטריים העשויים להסביר הטיה אפשרית.

הביקורת על אבחון הוגנות באמצעות חוסר-הטיה בניובי

טעון אחר של המחברים הוא שאפלילית הקבוצות הchallenge גלווה בעצם הניסיון לבנות מודדים המשמשים למימון (כגון ציון הבדיקה הפסיכומטרית) שיימשו את היעד של חוסר-הטיה בניובי. הנמקת מتابסת על הטענה שמהתנאי המקובל לחוסר-הטיה בניובי נובע בהכרח, שהפער בין הקבוצות בממוצע המדי גדול מהפער בינו לבין ממוצעי ההצלחה בليمודים. במלים אחרות, העדר הטיה בניובי אינו מתישב עם התנאי המוגדר על ידם כהעדר הטיה בברירה הכרחית.

המחברים מtabסים על הגדירה לתנאי (Cleary, 1968) הדורשת שקווי הרגרסיה לנינוי הקריטריון בשתי הקבוצות יתלכדו לקו רגרסיה אחד (ראה הדגמה בציור 1א). לאחר שטיבעו של קו רגרסיה טיפוסי קטן מ- 1, דרישת זו לא תוכל להתמלא אלא כאשר הפער בממוצעי החזאי בשתי הקבוצות גדול מהפער בממוצע הקריטריון בשתיו. זו אמנס התוצאה אותה גוזרים המחברים (שוון, 3, ע' 437, במאמרם מותני חוסר-הטיה בניובי:

$$(M_{x1} - M_{x2}) = (M_{y1} - M_{y2}) / r_{xy} \quad (1)$$

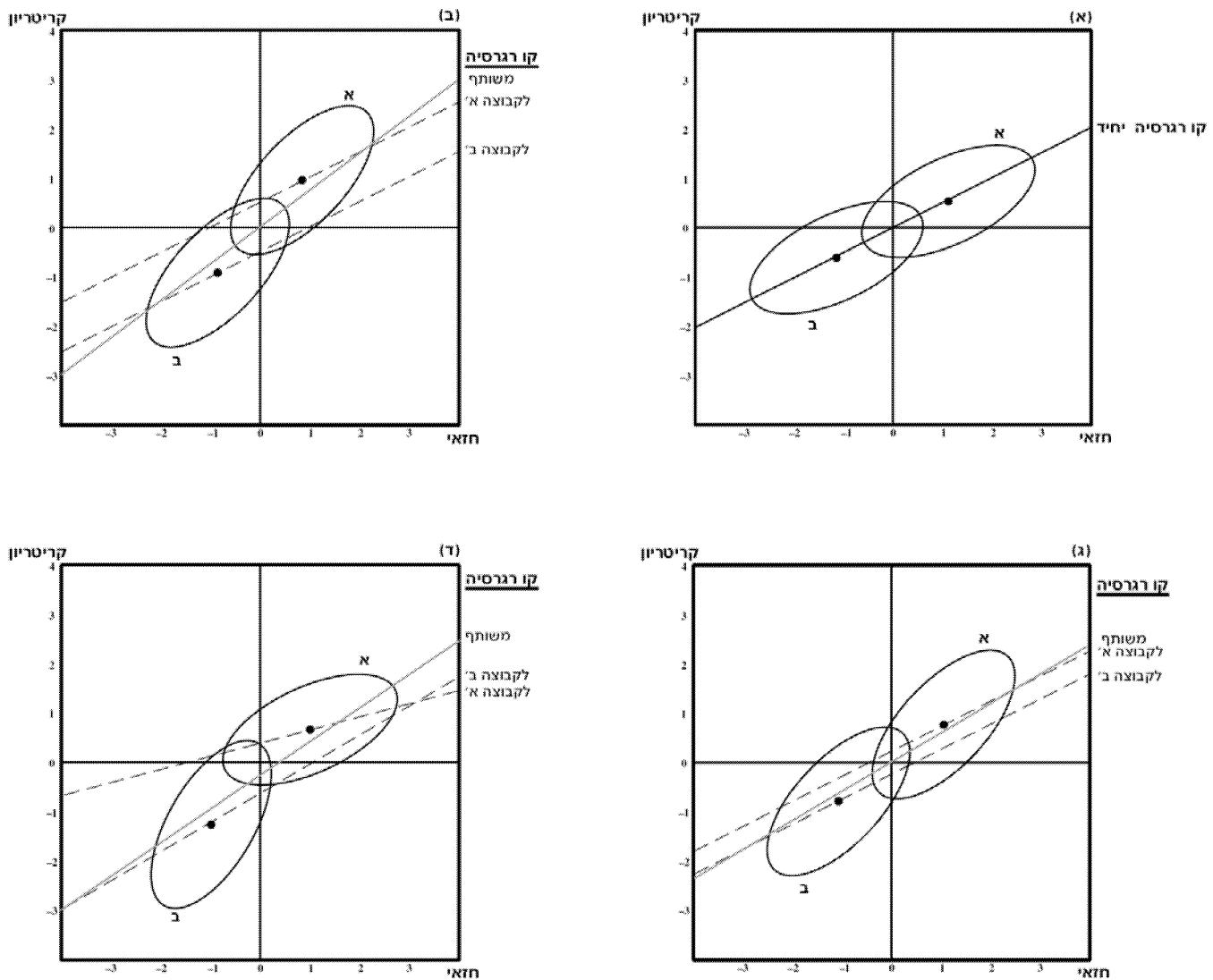
(כאשר M_{xi} מצין ממוצע חזאי בקבוצה i , M_{yi} מצין ממוצע קריטריון בקבוצה i , r_{xy} מצין את מקדם המתאם הליניארי בין החזאי לקריטריון).

הבעיה אינה בוגירה אלא במסקנה שהמחברים מסיקים ממנה.

ראשית, אם נסכים שהטיה מוגנית מוצמת להשפעת גורמים תרבותיים, חברתיים ואחרים שאינם הבדלי כושר, סביר להגיד רשות-הטיה לעל פי Birnbaum, 1979, 1981, על פי (בהתוח של הוגנות שכר) כמצב בו לשונות הקבוצה אין השפעה ישירה לא על החזאי ולא על הקריטריון, ועל כן כל הבדל נצפה בין הקבוצות נובע אך ורק מבדלים בכושר. אלא שהשפעת הכושר על המדי המשמש לקריטריון עשויה להיות קטנה מהשפעתו על החזאי (למשל, כאשר הקריטריון מושפע יותר מגורמים אחרים, כגון יכולת להකידיש זמן רב

ללימודים), כך שהפער בין הקבוצות במדד זה יהיה קטן מהפער המקורי על החזאי. מה שהמחברים מכנים "הтиיה בברירה הכרחית" הוא אפוא תוצאה אפשרית של מצב של חוסר-הтиיה על-פי הגדרה שדומה כי קשה לחלק עליה.²

שנית, המוחברים מניחים שמן הממצא של העדר אינדיקציה להтиיה שיטתייה בניבוי ציוני הלימודים באוניברסיטה (ראה קנט-כהן, 2001) ניתן להסיק שדרישתה של Cleary מתקיימת במקריםים. אלא שלמעשה מתיחסים רוב הממצאים לאבן בוחן מתחכמת יותר להтиיה בניבוי.



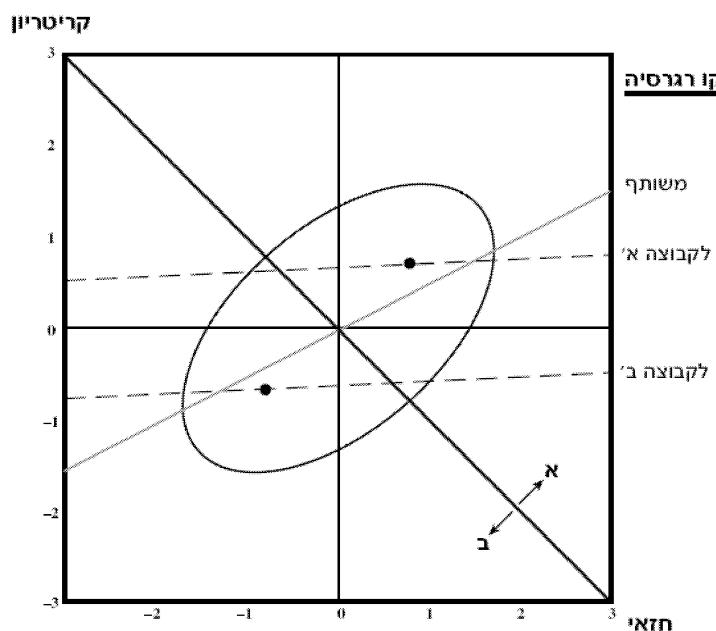
ציור 1 :

דוגמאות למבנים אשר בכל אחד מהם יש שתי קבוצות, אי ו-ב', השונות זו מזו בממצאי חזאי ובממצאי קרייטריון. באנל' אי מוגם מצב בו קווי הרגרסיה לשתי הקבוצות מתלכדים לקו חיד. באנל' ב' מוגם מצב בו לכל אחת מהקבוצות קו רגסיה מסוימת, כאשר הקווים שווים זה לזה במקדמי השיפוע אך שונים בנקודות החיתוך עם האורדיינטה (והקו המשותף אינו אלא הממוצע מעבר לכל האוכלוסייה). באנל' ג' מוגם מצב דומה, פרט לכך שהפער בין הקבוצות בקרייטריון קטן מהפער המקורי בחזאי, העשייה היוצר כאשר לכושר השפעה קטנה יותר על החזאי. באנל' ד' מוגם מצב בו מקדמי השיפוע שונים.

2. ניתן לטעון שהגדרה זו עדין צרה מדי. המושג הטיה מתייחס למעשה לחישום מהאפשרות שמשתנה הקבוצה ישפיע על החזאי ישירות, לא רק באופן עקיף דרך השפעתו על כושר אקדמי. מובן זה אין הטיה בהשפעה של משתנה הקבוצה על המודד בו משתמשים בקרייטריון דרך גורם שאינו קשור לכושר אקדמי. במקרה שיש השפעה כזו צריך להתעורר ספק אם הדרך הנכונה לתיקוף החזאי היא ע"י יכלהתו לנבא אותה מודד. מכל מקום, יש להביא זאת בחשבון בדיון באפשרות שהקרייטריון עצמו מוטה. בהמשך נדונה האפשרות שזה אמנס המצביע במקורה הנזון ע"י המוחברים.

לפי המודל של Birnbaum (1979, 1981) המתבסס על ניתוח נתיבים, הגדרת חוסר-ההטיה מתייחסת למשתנה תיאורטי, כושר. לאחר שימושה זה אינו מديد ישירות, אבחונה של חוסר-ההטיה על-פי משתנים נכפים מותנה בהנחות. כפי שנבע ממודל זה (ראה Linn; קנת-כהן, 2001), דרישת השם Cleary מתאפשרת רק בתנאים שאינם סבירים ביוטר (כשהקבוצות שוות בכושר, או כשהcoresר אינו משפיע על הקритריון, או כשהחוזאי המשמש למילוי תקופת מושלמת). תחת הנחות סבירות יותר לא ניתן לשולק קיומה של חוסר-ההטיה גם כשקווי הרגרסיה אינם מתלכדים.

לשם המחשה נتبונן במקרה היפותטי קיצוני בו הקבוצות מוגדרות אך ורק על-פי היישגיהן בחוזאי ובkritirion. נניח שבאוכלוסייה כולה ההתפלגות המשותפת של החוזאי והkritirion היא בינורמלית עם מקדם מתאים 0.5 (ראה ציור 2). שיפוע קו הרגרסיה של y על x בציוני תקן שווה לפחות אף הוא ל-0.5. נחלק את האוכלוסייה לשתיים עיי' האלכסון השילילי, ככלומר למחציתה הגבוהה בשני המשתנים ($y > x$) ולמחציתה הנמוכה בשני המשתנים ($y < x$). לאחר שהקבוצות מוגדרות אך ורק על-פי היישגיהן, אין כל סיבה להניח כאן קיומה של הטיה לשתי. כפי שניתן לראות, ממוצעי הקבוצות שנוצרו עיי' חלוקה זו אינם נמצאים על קו הרגרסיה המשותף. לכל קבוצה קו רגסיה משלה. לקווי הרגרסיה התוך-קבוצתיים מקדמי שיפוע שונים אך מקדמי חיתוך (אינטרפטיטים) שונים. במקרה זה אי אפשר כמובן לגזור את הפרש בין ממוצעי הקבוצות בkritirion בעורת שווין 1 שליל.



ציור 2 :

הדגמה לכך שאין התלכדות של קווי הרגרסיה התוך-קבוצתיים במצב בו על-פי הגדרת הקבוצות לא קיימת הטיה (ראה הסבר בגוף הטקסט). אין כל חפיפה בין תחומי הציונים של הקבוצות, והאלכסון השילילי מייצג את הגבול ביניהם.

אם כן, להסקה סבירה שאין הטיה בניבוי אין צורך בדרישתה המפליגת של Cleary, ועל כן היא זכתה לביקורת (ראה למשל Darlington, 1971; Linn, 1984) למروת שיש אליה התייחסות רבה בספרות (למשל, Maxwell & Arvey, 1993 ואוטם מצטטים המחברים). המלצת המקובלות לצורך בניית כלי מיל פסיקומטריים היא לעשות שימוש בהגדירה אופרציוונלית מתוקנת המקובלת על הזרים המרכזיים בתחום, הדוחה את השערת חוסר-ההטיה רק כמשמעותיים תנאי גבול מסוימים הן לגבי הרגרסיה של kritirion על

החזאי והן לגבי הרגרסיה של החזאי על הקרייטריוון. כך נעשה גם בبنית הבדיקה הפסיכומטרית לאוניברסיטאות בארץ (ראה למשל קנת-כהן, 2001, ע' 28-30).

הגדרה מתוקנת זו אינה שוללת בהכרח חוסר-הטיה במצב בו נמצא כי קבוצה חזקה וקבוצה חלשה נבדלות זו מזו במקדמי החיתוך של קווי הרגרסיה (ראה דוגמאות בציורים 1ב ו-1ג). היא אף אינה שוללת בהכרח חוסר-הטיה במצב בו מקדמי השיפוע של הרגרסיה אינם שוים (ראה דוגמה בציור 1ד). בהחלט ייתכן, למשל, שההتوزואה המודגמת בציור 1ג תיווצר כאשר יש הבדל אמיתי בקשר בין הקבוצות, אולם השפעת הקשר על החזאי גדולה מהשפעתו על הקרייטריוון. כמודגם בציורים 1ב ו-1ד, ייתכן שתיקיימש שוויון בעיריות בין הקבוצות בחזאי ובקריטריוון (קרי, "חוסר הטיה בברירה הרכיתית") אף כאשר לא ניתן להסיק על חוסר-הטיה בניבוי. ובכן, מסתבר שהמחברים משתמשים במקרה פרטיאי (התלכדות קווי ורגסיה) להוכחת טענה כללית (שהעדר הטיה בניבוי גורר את שוויון 1). מובן אףו מדוע הטענה שהוכחה מבקשת אינה בהכרח מתקינה.

יתר על כן, בנגד למושג תוקף הניבוי המתיחס לחיזוי ההצלחה בלימודים כשלעצמה, המושג הנומיינלי של הטיה מתיחס ל*אי-דיוק* שיטתי באומדן של המשנה המתוונ, קרי כושר אקדמי. אלא שהישגים בלימודים מושפעים לא רק מכשור אקדמי. ניתן להסיק אפוא על קיום הטיה של החזאי לרעת קבוצה מסוימת מאומדן-חסר של ההישגים בלימודים של אותה קבוצה רק במידה שהם מודדים כושר אקדמי טוב יותר מן החזאי. לעומת זאת, כמשמעותה הקבוצה משפיע באופן ניכר על ההישגים בלימודים דרך גורם שאינו קשור לכושר אקדמי (למשל, הקלה דיפרנציאלית בציוניים) ספק אם הפער בין הקבוצות על פניה מוצע הציוניים באוניברסיטה מהווה תשומה מספקת להערכת קיומו וגודלו של אומדן-חסר (ראה הערה 2). במקרה זה דרישתה של Cleary רלבנטית עוד פחות.

מסתבר שחשש זה מוצק למדי, לפחות כשמדבר במשמעות הציוניים באוניברסיטאות בארץ. על כך בפרק הבא.

פירוש הממצאים האמפיריים

במחקר שנערך לגבי ציוני הקבלה בהם משתמשות האוניברסיטאות בארץ (קנת-כהן, 2001) לא נמצא במרבית מסלולי הלימוד עדות להטיה שיטתי בניבוי (ניבוי דיפרנציאלי). נמצאה בהם עדות להבדלים קטנים בלבד בין יכולתם של החזאים לחזות את מידת ההצלחה לבני הקבוצה החלשה ליכולתם לחזות את מידת הצלחתם של בני הקבוצה האחרת (תוקף דיפרנציאלי).

אלא שעיקר טיעונים של המוחברים לגבי הוגנות מערכת המילון באוניברסיטאות בארץ מתיחס להגדרתם שלהם. נמצא בשורה של מחקרים (בלר ובנו-שחר, 1983 ; קנת וארון, 1988 ; קנת-כהן, 1987 ; 2001 ; Zeidner, 1987) שהפער בציוני אוניברסיטה בין קבוצות חברותיות שונות קטן במידה ניכרת מהפער בציוני הקבלה. הדוגמה אותה מציגים המוחברים היא הנתון המוצג בקנת-כהן (2001), לפיו בעוד הפער המתוון בציוני אוניברסיטה בין קבוצת מיצב חברתי-כלכלי גבוהה לבין קבוצת מיצב נמוך הוא 0.18, הפער בציון הקבלה המשוקל הוא 0.47 (כלומר, פי 2.6). על-פי ניתוח המוחברים להגדרות הוגנות, זהה ראייה ניצחת להעדרה. אך

עказ, גם אם קיבל עקרונית גישה זו, לא ברור כלל שהראייה מרשימה במידה בה היא מוצגת ע"י המוחברים. להלן יידונו מספר חולשותבולטות בדינום בממצאים. המשותף להן הוא הערכת חסר לגורמים סטטיסטיים הפעילים לכיווץ של סטטיסטיים נצפים בקבוצת המתקבלים יחסית לpermatorim מתקבלים של אוכלוסיית המועמדים. זה אמרור לא רק באפקט המוכר והידוע של כיווץ מקדמי תקיפות (המכונה בדרך-כלל "קיצוץ תחום"), אלא גם בכיווץ פערים בין קבוצות.

בעיית הקרייטריוון. החולשה הראשונה היא בהתייחסות לקרייטריוון. טיב המיון נבחן על-פי ממוצע ציוני הלימודים באוניברסיטה מפני שהוא המדד האופרציאונלי הזמין לקרייטריוון על-פי הגדרתו הנוומינלית, דהיינו לאותו צירוף משתנים הקובע את טיב הלמידה של הפרט במוסד להשכלה גבוהה (ఈ קושי הלמידה האובייקטיבי נתון). אך אף שמדובר הציונים נקרא בקיצור "קרייטריוון", אין הוא בהכרח משקף בצורה הטובה ביותר את מה שנועדה מדיית הקשר האקדמי לנבأ. יתרה מזו, אין בכוחו לכמת את סך התוצאה הנתרמת ע"י מיון מוצלח.

נתחיל דוקא בנקודת האחראונה. הממצאים אוטם בוחרים המחברים להציג נוגעים לכל המתקבלים מעבר לחוגים (ולמוסדות). פירוש ממצאים כאלה היה פשוט לו הייתה הקבלה מתבצעת על-פי סף אחד, כזה הקיים במוסד לימודים בעל מסלול יחיד. אלא שבכל אוניברסיטה יש מסלולים רבים, וציוון הקבלה משפיע מiad על התהליך בו מתמיינים המועמדים בין המסלולים לפי רמת יכולתם החזوية. זו כמובן אחת הפונקציות שלשלמן מקיימות האוניברסיטאות את מערכת הקבלה על פי חזאי הצלחה בלימודים, ולמעשה הפונקציה העיקרית במצב הקיים בו רובם הגדול של המועמדים נקלט בסופו של דבר באחד ממוסדות ההשכלה הגבוהה. למehrבה הצעיר, את האפקטיביות של מערכת הקבלה על תהליך ההתחמיינות לא ניתן למדוד ישירות, משום שאין כל ציוון המשותף לכל המסלולים, קל וחומר לכל המוסדות. טיב המיון לא נמדד אפוא על פי תרומתו לתהליך זה אלא על פי אומדן תקופות הנушה באמצעות חישובי מתאימים בתוך מסלולים,(Clomar על-פי שונות בציוני החזאי שלא שימושה בפועל לתהליך ההתחמיינות.³

יש לכך השלה גם על מדיית הפער בין קבוצות במדד המשמש לקרייטריוון. בשל האפקט הסטטיסטי הידוע בשם קיצוץ תחום, עצם המיון למסלול לימודי למוד על פי ספי קבלה שונים גורם לכך שהפער בין הקבוצות בתוך מסלול יהיה נמוך יותר מהפער ביןינו בכל המתקבלים לאוניברסיטאות (לא כל שכן בכלל המועמדים) אילו כולם היו לומדים באותו מסלול. לדוגמה, כאשר הפער בכלל המועמדים שווה ל-1.0, הפער בכלל המתקבלים (בנחה שהאוכולוסיות מתפלגות נורמלית וסף הקבלה הנמוך ביותר הוא ציון תקן 0.1-0) שווה ל-0.79, והפער בחוג סלקטיבי בו סף הקבלה הוא ציון תקן 0.5 שווה ל-0.51 (משמעות, בקירוב פי שניים קטן יותר מאשר בכלל המועמדים). בחוגים פחות סלקטיביים התמונה מורכבת עוד יותר. לעומת זאת, בשל סף הקבלה הנמוך יחסית קיצוץ התחום אמרה להיות מתון יותר. אלא ש למעשה רבים מהתלמידים בחוגים אלה לומדים בהם מפני שלא עמדו בסף הנדרש לחוג סלקטיבי יותר. עברו אלה קיצוץ התחום דוקא חריף יותר משום שבפועל הוא דו-צדדי. לדוגמה, בחוג בו סף הקבלה הוא ציון תקן 1.0- ובו נניח לשם פישוט כי כל התלמידים נדחו מחוגים שסף הקבלה שלהם הוא ציון תקן 0.5, הפער שווה ל-0.42 ; בחוג בו סף הקבלה נקבע לציון תקן 0.5- וגם בו כל התלמידים נדחו מחוגים שסף הקבלה שלהם הוא ציון תקן 0.5, הפער שווה 0.28 (משמעות, כמעט פי ארבעה קטן יותר מאשר בכלל המועמדים). במחקרה ערוכה קנת-כהן תיקונים ל垦יצוץ תחום עברו מקדם המתאים המשמש למדיית התקופות, אך ספק אם הם מכוונים בהכרח את כל השפעתו על ציון הפער בין הקבוצות, כאמור בהמשך.

.3. נדמה לעצמנו, למשל, מצב בו בסופו של תהליך ההתחמיינות לומד תלמיד אחד בכל מסלול. אם יש מתאם חיובי גבוה בין כושרים האקדמיים של התלמידים לבין הדרישות האקדמיות של המסלולים, הרי שהברירה על-פי החזאי אפקטיבית. בכך אין כל שיקוף בניתוחי התקופות המקובלים, מאחר שאלה משתמשים בשוניות הציונים בתוך מסלולים, שבמקרה זה הינה אפסית.

וחזרה לנקודת הראשונה. הערכת הפער בין הקבוצות בקריטריון מסתובכת עוד יותר בשל אופן המדידה של הצלחה בלימודים. זו מבוססת על מיצוע ציוניים שניתנו ע"י מורים שונים, בקורסים שונים ולפעמים גם במסלולים שונים. התפלגות ציוני האוניברסיטה של המתקבלים לא מאד שונות בין הקורסים ומסלול הלימוד חurf השוני בחומר הלימוד ובهرכוב התלמידים (ראה, למשל, Elliott & Strenta, 1988). אחת הסיבות לכך היא שמורים נוטים להוריד את רמת ההוראה ולהקל בציינון ככל שיכולתם המומצת של תלמידיהם נמוכה יותר (ראה, למשל, Goldman & Hewitt, 1975; Thorndike & Hagen, 1961).

תלמידים בעלי יכולת נמוכה ובינונית עשויים אףו לקבל ציונים גבוהים למדיה שהם לומדים יחד עם תלמידים רבים אחרים שיכולתם בינונית.

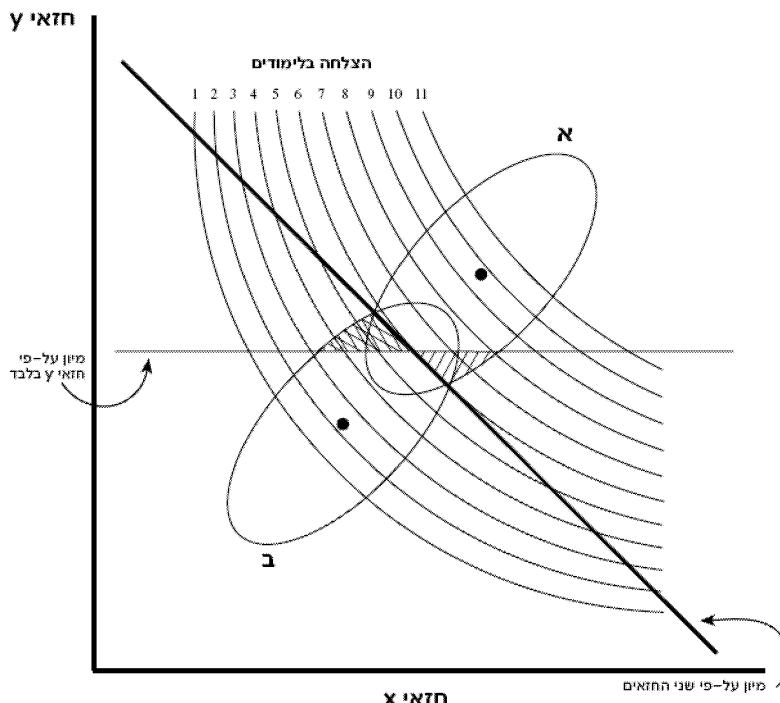
אין בכך רע בהכרח, שכן ייעודם של הציונים הוא בעיקר ליצור מדרג הישגים בקורס הנtanון ובבסיס למיון לתואר שני במסלול הנtanון. הבעיה מתעוררת כאשר ציונים משמשים להערכת המידה בה החזאי מודד כושר אקדמי ומגaba את התפקיד מהלימודים המהווה את הקריטריון הנומינלי, קרי הישגים במדידה בתנאי קושי קבוע. מבחינה זו יש כאן עיוות ברור – אומדן-יתר של הקריטריון. סביר שתלמידים שהם חברי הקבוצה החלשה נהנים ממוצע מעויות זה יותר מ学生们 שהם חברי הקבוצה חזקה, למרות שישוורם התלמידים מהקבוצה החלשה גבוההיחס בחולgas הפחות סלקטיביים (25.1% בחולgas הפחות סלקטיביים לעומת 15.2% בחולgas הסלקטיביים, ראה קנת-כהן, 2001, לוח 4, ע' 27), שבורובם כל יותר לקבל ציונים גבוהים. מאחר שהטיה בין-קבוצתית, במובנה החשוב מעשית, אינה מתייחסת לדיקן אומדן הרמה המומצת של קבוצה בקריטריון הנמדד אלא לדיקן אומדן הרמה המומצת שלא בכושר האקדמי ובתפקיד מהלימודים, יש בתופעה זו כדי להעלות ספק לגבי השימוש לצורך אבחון ההטיה בפער בין הקבוצות ממוצע הציוניים.

התופעה גורמת כמובן להפחחת הפער בקריטריון כאשר ממוצעים אותו מעבר לחולgas (השוואה Ramist, Lewis & McCamley, 1999) וכן לפער נמוך יחסית בחולgas הפחות סלקטיביים. ואולם, עיון נוסף נושא במצאי קנת-כהן מלמד שהפער בציוני האוניברסיטה בקרב המתקבלים משתמש הקבוצות בחולgas הפחות סלקטיביים קטן פי 2.4 מהפער המקביל בחולgas הסלקטיביים (קנת-כהן, 2001, לוח 6, ע' 36). ההשווואה בחולgas הפחות סלקטיביים לרבענית פחות כאמור, הוא בשל בעיטת הקריטריון בהם,/non מפני שיש בהם ייצוג יתר לתלמידים בעלי יכולת בינונית והן מפני שהאפקט של הברירה בהם מוקוץ ע"י אפקט מנוגד – של ריבוי תלמידים שנדרחו מחולgas סלקטיביים יותר. כאשר בוחנים את הפערים בחולgas הסלקטיביים, בהם ההשווואה לרבענית יותר, מנת הפערים בין הקבוצות (קרי הפער במדד המשמש למין מחולק בפער ממוצע הציוניים בלבד) היא 1.97 (בשיעור – 3.42 בחולgas הפחות סלקטיביים) – כלל לא בלתי-צפוי לאור סדר-הגודל של קיצוץ התחומים, הנחתת המתאימים (ראה לעיל בפרק "ההגדרה") וגורמים נוספים הנדרדים להלן.

בעיטת מדידת ההטיה. חולשה שנייה היא בחירותם של המחברים להתייחס לנתן הפערים הגדולה ביותר מותקן כל אלה שחושו ע"י קנת-כהן מבלי לתת את הדעת לסיבה להבדלים במנות הפערים. הפער בציון המשוקל (0.47) גובה מהפער בכל אחד ממרכיביו (0.41 בציון הבחינה הפסיכומטרית ; 0.40 ממוצע הבגרות), ויחס דומה מתקיים בין הפער בציון הכלול של הבחינה הפסיכומטרית לבין הפערים בשניים משולשת מרכיביה. בכך יש הסבר פשוט בגורם סטטיסטי חף מכל הטיה חברתית.

פער בציון מצורף כלשהו צפוי להיות בדרכ-כל גודל יותר מכל אחד מהפערים במרכיבים. כאשר המרכיבים בלתי- תלויים, הפער בציון המצורף שווה לשורש הריבועי של סכום ריבועי הפערים (ראה ע' 239 ב-Green & Swets, 1966). לשם המראה, פער זה שווה $\sqrt{2}$ כאשר כל אחד משני הפערים שווה ל- 1. נקל להוכיח (ראה נספח 1) כי גם כאשר המרכיבים אינם בלתי- תלויים, סביר מאד שהפער בציון המצורף יהיה

גדול מכל אחד מהפערים במרקביים. כאשר, למשל, הציון המצויר הוא סכום שני מרכיבים שמקדם המתאים ביניהם שווה בין הקבוצות, ובאשר השווניות שווות בין הקבוצות ובין המרכיבים, תוצאה זו צפואה כל עוד המכפלה $(z+1)/z$ קטנה ממנה בין סכום הפערים לפער הגadol מביניהם. לדוגמה, כשהפערים שוויים המנה האחורה שווה ל-2, ומאחר שהמכפלה קטנה מ-2 (אלא אם כן המתאים מלא) התנאי מתקיים בוודאות. סביר יותר, אם כן, שהוא מתקיים במקרה של הציון המשוקל המשמש למילון לאוניברסיטאות, שכן הפערים דומים במידה מסוימת והמתאים ביניהם הוא כ-0.5 (ראה קנט-כהן, ברונר ואורן, 1999). אם כן, מנת הפערים בציון המשוקל צפואה להיות גבוהה יותר מאשר מנת הפערים בכל אחד מרכיביו, מפני שהמונה של המנה גדוֹל יותר בציון המשוקל בשל גורם סטטיסטי טהור. דוגמה זו, ממש כדוגמה אפקט המהימניות הנדון לעיל (בסוף פרק "ההגדרה"), ממחישה אף היא את הביעתיות שבഗדרת הוגנות על-פי שוויון בין פערים. מאחר שגודל הפער בחזאי תלוי במספר מרכיביו של אותו חזאי, ניתן להקשוט למשל אם יעלה על הדעת שמיון הוגן כאשר הוא מבוסס על סכימת מרכיבים שככל אחד מהם עדיין הוגן מספיק. השלכתה המעשית של בעיה זו לפירוש הממצאים האמפיריים של קנט-כהן מחייבת את מסקנת המחברים: גם אילו קיבלונו את הנחתם הבסיסית שמנת פערים ניכרת משקפת חריגה ממילון הוגן, אין הצדקה להבלת המנה הגדולה ביותר ללא התייחסות לסיבה הסטטיסטית לכך שהיא גדוֹלה מהאחרות. בעיית השפעת גורמי התקפות. אך בזאת לא תמו הגורמים הסטטיסטיים המשפיעים על גודל המנה, שכן זה תלוי גם בגודל המכנה (זהיינו, הפער בקריטריון). כאן פועל בין היתר גורם סטטיסטי אחר – מספר החזאים לפיהם מתבצעת הברירה. כאשר משווהת הקבלה היא צירוף ליניארי של חזאים, הפער בקритריון צפוי להתכווץ יחסית למצב בו הקבלה היא על-פי אחד החזאים בלבד. הסיבה לכך מוכבת במקצת: כאשר לשווהת מיון הכוולת ציון בחזאי ע בלבד מושפעים ציון בחזאי א (ראה ציור 3 מטאפרשת קבלה של מועמדים שציוניים בחזאי ע אינם גבוהים אך ציוניים בחזאי א גבוהים מספיק כדי לפצות על כך (ראה השיטה המקווקו בציור בקווים מלוכסנים), זאת על חשבו מועמדים אחרים שבגובה ציוניים בחזאי ע או די כדי לפצות על ציוניים הנמוכים בחזאי א (ראה השיטה המקווקו בתבנית רשות). לא במפתח, ציון הקריטריון המוצע של הראשונים גבוה מזה של האחרונים (להנאה מפורת יותר ראה נבוּן, בן-שחר וצ'לוב, 2001). חשוב יותר לעניינו: מחישובים שערכנו באמצעות סימולציה מונטה-קרלו (ראה תוצאות בלוח 1) מסתבר גם ש בקרב הראשונים הפער בין הקבוצות בקריטריון נמוך בשיעור ניכר מאשר בקרב האחרון, ועל-כן הפער בקריטריון מעבר לכל המתקבלים נמוך יותר מאשר אילו נכלך רק חזאי יחיד במשווהת המילון. הסיבה לכך היא שתופעת חזאי לשווהת המילון משפיעה יותר על קבוצה בה שיעור המתקבלים מלבתילה קטן מאשר על קבוצה בה שיעור המתקבלים גבוה ממילא ועל-כן רובם אינם מושפעים משיוני משווהת המילון.



ציור 3 :

דוגמה להשפעת תוספת חזאי למשוואת המיון על הפער בממוצעי ציוני הקרייטריון בין קבוצות אוכלוסייה יחסית למשוואת מיון המבוססת על החזאי הקודם בלבד. האליפסות מייצגות היטלים של התפלגיות דו-משתניות של שני חזאים, x ו- y , כל אחת עבור קבוצה אוכלוסייה שונה. הקו האופקי מייצג סך מימון על פי חזאי y בלבד. הקו האלכסוני מייצג סך מימון על פי צירוףلينארי של שני חזאים. הקווים הקמורים כלפי הראשית מייצגים עקומות שווות-הצלחה, כאשר הצלחה גבוהה יותר ככל שגדל המרחק מהראשית. (ראה הסבר בגוף הטקסט).

ЛОח 1 :

הפער הצפוי בממוצעי ציון הקרייטריון בקרב המתקבלים על-פי משוואות מימון המבוססות על חזאי בודד y , או צירוף חזאים x ו- y , כאשר תוקפו האמיתי של חזאי y הוא 0.45, ותוקפו האמיתי של חזאי x הוא 0.45, המתאים בין החזאים הוא 0.51, ממוצע הקבוצה הגבוהה הוא 0.5, ממוצע הקבוצה הנמוכה הוא 0.5, בכל המשתנים, והשינויות הן 1.0 בכל המשתנים ושתי הקבוצות, עבור שלושה יחסית סלקציה שונים, 1.0, 0.5, 0.33.

סף הקבלה	יחס הסלקציה	מיון על-פי	בציוון הקרייטריון	פער צפוי בקרב המתקבלים
0.67	0.67 (67% מתקבלים)	חזאי בודד	0.8185	חזאי בודד
-0.5		ציון מצורף	0.7355	ציון מצורף
0	0.50 (50% מתקבלים)	חזאי בודד	0.7635	חזאי בודד
0.5	0.33 (33% מתקבלים)	חזאי בודד	0.6485	ציון מצורף
0.5	0.33 (33% מתקבלים)	חזאי בודד	0.7170	חזאי בודד
		ציון מצורף	0.5830	ציון מצורף

הערות :

התוצאות הן ממוצעי שתי הרצאות מונטה-קרלו, שבכל אחת מהן N שווה ל-100,000. הפער מחושב ע"י ממוצע שתי שוויונות המדגם בשקלול שווה. בחישוב חלופי, ע"י שימוש בשנוות הקבוצה הגבוהה בלבד, התקבלו תוצאות דומות למדי.

יתר על כן, הגדלת מספר החזאים אינה הגורם היחיד הפועל לצמצום הפער בקריטריון. במקרה שתתקפות שני החזאים גדולות יותר מזו המונחת בלוח 1, 0.60 במקומ 0.45, מתקבלים בסימולציה מונטה-קרלו (עבור סף קבלה 0) פערים קטנים יותר – 0.71 לחזאי בודד ו- 0.64 לחזון מצורף.

ובכן, החולשה השלישייה בטיעוני המחברים דומה במקורה לשנייה – חוסר תשומת לב מספקת לכך שוגדי הפערים תלויים בראש וראשונה בגורמים סטטיסטיים. הניתוח האחרון מלמד, שמנת הפערים גדלה לא רק ככל שהחזאים מהימנים יותר, אלא גם ככל שהם מרוביים יותר ותקפים יותר. הגדרת הוגנות על-פי שוויון בין פערים נראית תמורה שבעתים כאשר מסתבר שהפער בקריטריון מצטמצם ככל שהחבריה מתבססת על מידע רב ותקף יותר. מאחר שלא מתקבל על הדעת שיפור בטיב המיון מפחית בהכרח את מידת ההוגנות,סביר יותר שהבעיה היא ב"מידדת" ההוגנות.

בעיית אמידת הפער בין קבוצות המועמדים. לבסוף, האם הממצאים הנדונים שחושו כМОן מצווני האוניברסיטה של המתקלבים יכולםאמין להיחשב לאומדן חסר-הטיה לוגדל הפער בקריטריון בין קבוצות המועמדים (כלומר, זה שהוא נמצא לו התקבלו כל המועמדים) אליו מתיחסת הגדרתם של המחברים להטיה חברתית?

בכל שיטת ברייה המבוססת על חזאי הצלחה, נדחים אלה שהחזאי מסמן כבעל פוטנציאל נמוך יותר להצלחה. שיעורם של אלה גבוה יותר בקבוצת הצלחה לפי הגדרה. משום כך סביר שבקבוצת הצלחה ממוצע ציוני האוניברסיטה של המתקלבים יהיה גבוה במידה ניכרת ממוצע הציונים של כלל המועמדים אילו התקבלו כולם. סביר שהמיון ישפייע פחות על ממוצע הציונים בקבוצת החזקה, כיון שמתוכה נדחין פחות מועמדים. אם כן, הפער בין הממוצעים בקבוצת המתקלבים יהיה נמוך מהפער שהיה נצפה לו התקבלו כל המועמדים (והשווה 1997 Lewis & Willingham).

אפקט זה גדול יותר מאשר הפיזור בקבוצת החזקה קטן מהפיזור בקבוצת הצלחה. הדוגמה היפותטית הבאה אולי תסייע להמחשת הנקודה: נניח שממוצע הגובה של קבוצת נערים לפני ג'יסס הוא 175 ס"מ, ואילו ממוצע הגובה של קבוצת ילדי בית-ספר תיכון מכל שיש השכבות הוא 155 ס"מ. הבה נניח עוד שההפרש, 20 ס"מ, שקול ל- 1.5 בציוני תקן. נניח שלמתקן ספורט מסוים מרשימים להיכנס רק למי שגובהו 150 ס"מ ומעלה. היש ספק שההפרש בין ממוצעי הגובה של חברי שתי הקבוצות שהורשו להיכנס למתקן יהיה קטן בהרבה – 20? סביר אפוא שם תימדד בתוך המתקן רמת הביצוע בפועלות כלשהי שיש לה מתאפס חיבוי עם גובה, הפער המתווך בממוצע שתי הקבוצות יהיה נמוך במידה ניכרת מ- 1.5.

הנוסחה לאמידת הפער בכל המועמדים בה השתמשה קנות-כהן (שפותחה מנוסחה שהוצאה ע"י גייקוב כהן, 1988, k. 24) שעלה לתשומנה מסטמכיים המחברים, אומדת אותו בעורת תיקו לקיצוץ תחום עבור מקדם המתאים. אלא שטיב האומדן באמצעות שיטה זו, מתחוכמת כאשר תהא, תלוי בהנחות. התיקו לקיצוץ תחום נעשה מעבר לקבוצות האוכלוסייה ובמובלע מניה פיזור שווה בתוך כל אחת מהן.

הנחה זו אינה תופסת בהכרח לגבי סולם הציונים בעל הטווח הנטוון (0-100) בו משתמשים במוסדות להשכלה גבוהה: לאחר שבד"כ נוטות התפלגיות הציוניים להתרכו סביב ערך אמצעי הנמצא בחציו העליון של הטווח (לדוגמה, ממוצע ציוניו שונה בחוג לפסיקולוגיה באחת האוניברסיטאות בארץ בשנים תשנ"ב-תשנ"ו הוא כ-88 וסטיית התקן 6.3), לא בלתי-סביר להניח שגם התקבלו כל המועמדים, התפלגות ציוני הקריטריון של המועמדים מקבוצת אוכלוסייה חלה יחסית הייתה מפוזרת יותר על-פני הטווח מאשר ההתפלגות המקבילה בקבוצת החזקה יותר. משום כך אומדן הפער על בסיס ההנחה של שוויון שוניות עלול להיות אומדן-חסר.

כדי להדגים זאת ערכנו ניתוח המבוסס על סימולציה של תוצאותיו של תהליך מיוון על-פי פרמטרים ידועים. האלגוריתם הבסיסי בניתו מתוכנת לחשב השפעת קיצוצים שונים על הממצאים והשונויות של מתקנים משתי קבוצות כאשר כל אחת מאוכלוסיות המועמדים מתפלגת בינורמלית עם פרמטרים נטוניים. לאחר שבמציאות ידועים דוקוא הסטטיסטיים של נתוני הצלחה בלמידים של המתקנים בלבד, מנסה הסימולציה באתר מודל הנוטן קירוב אופטימלי לסטטיסטיים אלה בעזרת הפעלת האלגוריתם על אלף צירופים של ערכי אפשריים של הפרמטרים הבלטי-ידועים (ממוצעים ושונויות של הצלחה פונצייאלית בלימודים בשתי אוכלוסיות המועמדים, מקדמי התקפות בשתייה וסף קבלה) והתכניות לעבר הצורך המספק התאמה מרבית.

יצרנו שני בסיסי נתונים פיקטיביים בהם ציוני הצלחה בלמידים מתפלגים נורמלית בשתי קבוצות אוכלוסייה סביר ממוצעים שונים (0.4- ו- 0.4), בסיס נתונים אחד בו סטיות התקן שווות ושני בו סטיית התקן של הקבוצה הדרישה גדולה פי שתים מסטיית התקן של הקבוצה החזקה. לאחר מכן השתמשנו באלגוריתם כדי לשחזר את הפרמטרים שאינם ידועים בתנאי המציאות. הופעל תהליך חיפוש ממוחשב אחר אותו צירוף ערכי הפרמטרים הבלטי-ידועים המספק את ההתאמה הטובה ביותר (במנוחה טעות RMS) לנוטנים. במקביל ביצענו ניתוח נוספת, זהה לזו ערכה קנת-כהן. תוצאות שני הניתוחים היו דומות למדי בתחום סטיות התקן בו סטיית התקן של הקבוצה הדרישה גדולה יותר התקבלו שמנת הפערים בהתאם. לעומת זאת, בתחום סטיות התקן של הקבוצה הדרישה גבוהה יותר התקבלו שמנת הפערים המתකבלת מהחישוב של קנת-כהן גבוהה ב- 51% (1.25 לעומת 1.89). מمنت הפערים שאפיינה את הנתונים אותם יצרנו (שוחזרה ע"י האלגוריתם בדיק נמרץ, כאשר הוא מדוחה על שיעור טעות מבוטל של 0.002. **יחסות טעות RMS**).

מעיוון בנתוני הצלחה בלמידים של המתקנים לא נראה היה שהנחה שוויון השונויות מופרת באופן בלט בחלק גדול מהمسلمולים. מכאן, ניתן ששוויון השונויות בקרב המתקנים נובע מקיוץ תחום דרמטי. לא ניתן אפוא לקבוע שאמנים חלק מمنت הפערים הגבוהה בנתוני קנת-כהן נובע מהנחה מוטעה של שוויון שונויות באוכלוסייה המועמדים. אך האפשרות בעינה עומדת, לאחר שהחישוב אותו ערכה מן הסתם אינו מביא בחשבון השפעות של אותם גורמים הנוגעים להתמיינות המתקנים שנדרשו לעיל, כגון סלקציה עצמית וקיוצי תחום מלמעלה. אם אכן הייתה הטעיה כזו, אין היא מעוררת את הממצאים שלשם ערכה קנת-כהן את החישוב, אך יש בה כדי להטיל צל על המסקנות אותן מסיקים גמליאל וקהאן מוגדל מנת הפערים שהם חישבו על סמך מצאה.

מאחר שהאלגוריתם המתוואר כאן מניה סף קבלה אחיד ואף הוא אינו מתחשב בגורמים שנמננו לעיל, אין הוא מתחכם די כדי לשחזר את הפרמטרים של כלל אוכלוסיית המועמדים אותן חישבה קנת-כהן. מכאן, כל אותן גורמים מעלים ספק כאמור לגבי תקופות המודל הסטטיסטי אותן החישוב בו השתמשה קנת-כהן. על-כן, לשם השוואת הפעלים את האלגוריתם על שני מוגדים מהווים חלק מנתוני קנת-כהן, וגם חישבנו בהם מנות הערים על פי שיטתה.

נלקחו ציוני הסכם (המורכבים משקלול הבחינה הפסיכומטרית וצינוי הבגרות) והצלחה בלמידים של כל תלמידי החוג לכלכלה באוניברסיטה העברית שהחלו ללימוד בשנים 1991 ו- 1992, ושל כל תלמידי הפקולטה למשפטים באוניברסיטת תל אביב שהחלו ללימוד בשנים אלה. הופעל תהליך חיפוש ממוחשב אחר אותו צירוף ערכי הפרמטרים הבלטי-ידועים המספק את ההתאמה הטובה ביותר (במנוחה טעות RMS) לנוטנים. בניתו זה התקבל שהאומדן למנת הפערים בין ממוצעי התפלגיות המועמדים משתי הקבוצות הוא 1.34 בכלכלה ו- 2.30 במשפטים. בניתו נוספת של נתונים אלה, זהה לזו ערכה קנת-כהן על כלל

הנתונים, התקבלו מנוגת הפערים הבאות : 2.24 בכלכלה ו- 1.88 במשפטים. לאחר שבאחד משני המדגמים אומדנה של קנט-כהן גבוהה כמעט פי שניים מזוה המתקבל מהאלגוריתם, לא ברור כלל שניית לסמוֹך עליו כאומדן תקף. יתר על כן, לאחר שהאלגוריתם לא הצליח להתכנס לאומדנים לדיקוק גבוהה (שיעור טעויות RMS של 20.73 יחידות בכלכלה ו- 26.33 יחידות במשפטים), נראה שלא מתקיימות לפחות חלק מההנחות המונחות הן בסיסו והן בסיס השיטה בה השתמשה קנט-כהן, למשל ליניאריות קווי הרגRESSED או בינוי רמליבות ההתפלגיות. מסקנה זו מעלה ספק נוספת באשר לטיב האומדן למנת הפערים עליו הסטמכו גמיליאל וקהאן במאמרם.

סיכום. ננו לעיל מספר גורמים פסיכומטריים ואחרים הופיעלים לכיווץ הפער בין קבוצות אוכלוסייה מוחזאים לקריטריון (בדומה לגורמים מעין אלה הופיעלים לכיווץ מקדם התקפות מאוכלוסיית המועמדים לקבוצת המתקבלים) - קיצוץ תחום, הנחתת מתאימים עקב מהימנות פחותה, הנחתה יתרה בחוגים לא סלקטיביים, הגדלת מספר החזאים במשוואת הברירה והגדלת התקפות. לבסוף, מנת הפערים בין הקבוצות צפואה להיות גדולה יותר מאשר בוחרים ממונה דזוקא בפער בציון משוקל של כמה גורמים ולא במדד ייחיד. בנוסף לכך, הtgtלו ספקות לגבי הדיק בاميידת הפער בקריטריון בין קבוצות המועמדים עליה מסתמכים המחברים. כפי שהוזג לעיל, סביר מאד שהשפעות המצרפית של כל הגורמים הללו מסבירה את רוב היפוי הנצפה בקבלת אוניברסיטאות או אפילו את כולם. הניסיון לייחסו בעיקר להטיה מובנית בהתעלם מכל אותם גורמים אינו נשען על כל ראייה נוספת. כל עוד לא סופקו ראיות כאלה, אין בסיס להעדיף את הטענה שנמצאה הטיה מובנית.

מראוי מקומ

- בלר, מ., ובן-שחר, ג. (1983). על הוגנות השימוש בבחינות הפסיכומטריות. *מגמות*, *ב*, 56-42.
- גמליאל, א., וקחאן, ס. (2004). חוסר הוגנות בקבלה לאוניברסיטאות: הגרסה המודרנית של משל "כbast הרש". *מגמות*, *מג*, 445-433.
- _nbrown, ד., בן-שחר, ג., וצלגוב, י. (2001). ההודעה על מות הפסיכומטרי הייתה מוקדמת מדי: תגובה ליגוב ואיילון (2000). מאמר שהופץ ע"י המחברים לרשות תפוצה מוגבלת (הופיע גם ב-2006 כדו"ח מס' 328. ירושלים: מרכז ארכי לבחינות ולהערכתה).
- סטולר, ר. (2005). דו"ח סטטיסטי למועדים: אוקטובר, דצמבר 2004 ופברואר 2005. (דו"ח מס' 322). ירושלים: מרכז ארכי לבחינות ולהערכתה.
- קנתק, ת., ואורון, כ. (1988). בדיקת ההוגנות התרבותית בשימוש במערכת המיון בשתי אוניברסיטאות (דו"ח מס' 78). ירושלים: מרכז ארכי לבחינות ולהערכתה.
- קנתק-כהן, ת. (2001). ניבוי דיפרנציאלי ותוקף דיפרנציאלי של מערכת המיון לאוניברסיטאות על-פי מיצב חברתי-כלכלי של המועדים (דו"ח מס' 285א). ירושלים: מרכז ארכי לבחינות ולהערכתה.
- קנתק-כהן, ת., ברוור, ש., ואורון, כ. (1999). ניתוח-על של תוקף ניבוי של מרכיבי מערכת המיון לאוניברסיטאות בישראל כלפי מידת ההצלחה בלימודים. *מגמות*, *מ*, 71-54.
- Alba, M., Butler, M., Dennis-Elmore, M., Johnson, W., MacDougall, S., Rudert, E., & Zieseniss, M. (2002). Beyond percentage plans: The challenge of Equal Opportunity in Higher Education (Staff Report). Washington, DC.: U.S. Commission on Civil Rights, Office of Civil Rights Evaluation, November 2002.
- Birnbaum, M.H. (1979). Procedures for detection and correction of salary inequity. In T.R. Pezzullo & B.F. Birtingham (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.
- Cahan, S., & Gamliel, E. (2001). Prediction bias and selection bias: An empirical analysis. *Applied Measurement in Education*, *14*, 109-123.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Cole, N.S. (2001). The new faces of fairness. *Journal of Educational Measurement*, *38*, 369-382.
- Cleary, T.A. (1968). Test bias: prediction of grades for Negro and White students in integrated colleges. *Journal of Educational Measurement*, *5*, 115-124.
- Darlington, R.D. (1971). Another look at "cultural fairness". *Journal of Educational Measurement*, *8*, 71-82.
- Elliott, R., & Strenta, A.C. (1988). Effects of improving the reliability of the GPA on prediction generally and on comparative predictions for gender and race particularly. *Journal of Educational Measurement*, *25*, 333-347.

Equal Employment Opportunity Commission. (1978). Uniform guidelines on employee selection procedures. Federal Register, 43, 38290-38315.

Goldman, R.D., & Hewitt, B.N. (1975). Adaptation-level as an explanation for differential standards in college grading. Journal of Educational Measurement, 12, 149-161.

Green, D.M., & Swets, J.A. (1966). Signal detection theory and psychophysics. New York: Wiley.

Guilford, J.P. (1965). Fundamental statistics in psychology and education (4th ed.). New York: McGraw-Hill (Originally published in 1942).

Hunter, J.E., Schmidt, F.L., & Rauschenberger, J.M. (1977). Fairness of psychological tests: Implications of four definitions for selection utility and minority hiring. Journal of Applied Psychology, 62, 245-260.

Lewis, C., & Willingham, W.W. (1997). The effects of sample restriction on gender differences. In W.W. Willingham & L.M Johnson (Eds), Supplement to Gender and Fair Assessment (Report 2). Princeton, NJ: Educational testing Service.

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, 1-14.

Linn, R.L. (1984). Selection bias: Multiple meanings. Journal of Educational Measurement, 21, 33-47.

Maxwell, S.E., & Arvey, R.D. (1993). The search for predictors with high validity and low adverse impact: Compatible or incompatible goals? Journal of Applied Psychology, 78, 433-477.

Peterson, N.S., & Novick, M.R. (1976). An evaluation of some models for culture-fair selection. Journal of Educational Measurement, 13, 3-29.

Ramist, L., Lewis,C., & McCamley,L. (1990). Implications of using freshman GPA as the criterion for the predictive validity of the SAT. In W.W. Willingham, C. Lewis & L. Ramist (Eds.), Predicting College grades: An analysis of institutional trends over two decades (pp. 253-288). Princeton, NJ: Educational Testing Service.

Thorndike, R.L., & Hagen, E. (1961). Measurement and evaluation in psychology and education (2nd ed.). New York: Wiley.

Willingham, W.W., & Cole, N.S. (1997). Gender and fair assessment. Mahwah, NJ: Erlbaum.

Zeidner, M. (1987). Test of the cultural bias hypothesis: Some Israeli findings. Journal of Applied Psychology, 72, 38-48.

נספח 1 : היחס בין הפער המצויר לפערים במרכיביו

יהיו x ו- y שני ציונים, ויהיו 1 ו- 2 שתי קבועות אוכולוסייה זרות. נניח לשם פישוט שהן שוות בגודלן. ממצוע יסומן $b - M$ ושותות תסומן $b - S^2$.

יהא d_x הפער המתוקן בין קבועות האוכולוסייה 1 ו- 2 בציון x

$$\cdot (M_{x1} - M_{x2}) / [0.5 \cdot (S_{x1}^2 + S_{x2}^2)]^{1/2}$$

יהא d_y הפער המתוקן בין קבועות האוכולוסייה 1 ו- 2 בציון y

$$\cdot (M_{y1} - M_{y2}) / [0.5 \cdot (S_{y1}^2 + S_{y2}^2)]^{1/2}$$

יהא הציון המצויר w שווה לסכום הציונים $x+y$,

ויהא d_w הפער המתוקן בין קבועות האוכולוסייה 1 ו- 2 בציון w

$$(M_{w1} - M_{w2}) / [0.5 \cdot (S_{w1}^2 + S_{w2}^2)]^{1/2}$$

השווה ל-

$$[(M_{x1} + M_{y1}) - (M_{x2} + M_{y2})] / [0.5 \cdot ((S_{x1}^2 + S_{y1}^2 + 2r_1 S_{x1} S_{y1}) + (S_{x2}^2 + S_{y2}^2 + 2r_2 S_{x2} S_{y2}))]^{1/2}$$

השווה מצדוי ל-

$$[(M_{x1} - M_{x2}) + (M_{y1} - M_{y2})] / [0.5 \cdot ((S_{x1}^2 + S_{y1}^2 + 2r_1 S_{x1} S_{y1}) + (S_{x2}^2 + S_{y2}^2 + 2r_2 S_{x2} S_{y2}))]^{1/2}$$

הבה נניח עתה, לשם פישוט, שווינו בין האוכולוסיות במקדמי המתאים בין x ל- y , $r=r_1=r_2$, ובשותויות,

$$S_y^2 = S_{y1}^2 = S_{y2}^2, S_x^2 = S_{x1}^2 = S_{x2}^2 \text{ וכמו כן } d_w^2 = S_w^2 = S_{w1}^2 = S_{w2}^2$$

$$[(M_{x1} - M_{x2}) + (M_{y1} - M_{y2})] / [0.5 \cdot 2 \cdot (S_x^2 + S_y^2 + 2r S_x S_y)]^{1/2}$$

השווה כMOVEN ל-

$$[(M_{x1} - M_{x2}) + (M_{y1} - M_{y2})] / [(S_x^2 + S_y^2 + 2r S_x S_y)]^{1/2}$$

עוד נניח, לשם פישוט, שווינו בין שותויות שני הציוןים, $S_x^2 = S_y^2$. במקרה זה, d_w^2 שווה ל-

$$[(M_{x1} - M_{x2}) + (M_{y1} - M_{y2})] / [2S^2(1+r)]^{1/2}$$

השווה מצדוי ל-

$$[(M_{x1} - M_{x2}) + (M_{y1} - M_{y2})] / [S \cdot \sqrt{2} \cdot \sqrt{1+r}]$$

נגיד עתה את α כיחס בין סכום הפרשי הממוצעים בשני הציוןים, $[(M_{x1} - M_{x2}) + (M_{y1} - M_{y2})]$, לבין הഫפרש הגדל מビיניהם (משמעות, $\alpha < 1$). נניח, בלי אובדן כלליות, שההפרש הגדל הוא $(M_{y1} - M_{y2})$, כלומר,

$$[(M_{x1} - M_{x2}) + (M_{y1} - M_{y2})] = \alpha \cdot (M_{y1} - M_{y2})$$

$$\alpha \cdot (M_{y1} - M_{y2}) / [S \cdot \sqrt{2} \cdot \sqrt{1+r}]$$

ועל-כן גם ל-

$$\alpha \cdot d_y / [\sqrt{2} \cdot \sqrt{1+r}]$$

מכאן, כאשר $\alpha < \sqrt{2} \cdot \sqrt{1+r}$ מתקיים אי-השוויון $d_y < d_w$.

마וחר שבנהנת שותויות α שווה גם ליחס בין סכום הפערים המתוקנים לפער הגדל, אפשר להגיד גם בתנאי מספיק להתקיימות אי-השוויון שליל, שיחס זה יהיה גדול מ- $\sqrt{2} \cdot \sqrt{1+r}$.

ההשלכה המעשית לעניינו היא, שבנהנת שותויות שותויות d_x ו- d_y שוות בקירוב, $d_w < d_y$ עבור כל ערך r . האפשרי כמתאים ב מבחנים מהסוג בו דן מאמר זה.

