

הגורמים המשפיעים
על אומדן מקדם תקפות החזאי
המשמש לבררת מועמדים
להשכלה גבוהה

דוד נבון
יואב כהן
שמואל ברונר

יוני 2008



דוח מרכז 347
ISBN:965-502-140-8

הגורמים המשפיעים על אומדן מקדם תקפות החזאי המשמש לבררת מועמדים להשכלה גבוהה

שמואל ברונר

יואב כהן

דוד נבון

תמצית

חישוב תקפותם של חזאי כושר ללימודים אקדמיים נעשה בדרך-כלל באמצעות מיצוע מקדמי מתאם מעבר ליחידות לימוד (חוגים, מוסדות, מחזורים) לאחר תיקונים לקיצוץ תחום, וזאת בהנחה שהשונות במקדמי מתאם נצפים נובעת בעיקר ממידת קיצוץ התחום. בניתוח המדווה כאן נבדקה תקפות ציון הסכם לחיזוי הצלחה בלימודי שנה א' באמצעות מקדמי מתאם מתוקנים ב-778 יחידות לימוד, שמוינו על-פי אופי הדיסציפלינה הנלמדת ומידת הסלקטיביות בקבלה. נמצא שהמקדמים במדעים המדויקים גבוהים במובהק מאשר במדעי הרוח והחברה, וכי הם גבוהים במיוחד במסלולים הסלקטיביים ביותר. האפקט של אופי הדיסציפלינה מיוחס למהימנות המדדים המשמשים כקריטריון. את אפקט הסלקטיביות ניתן להסביר בכך שבמסלולים בהם הסלקטיביות אינה גבוהה פוחתת רגישות הקריטריון להבחנה בקצה העליון של התפלגות הכושר, כנראה בשל רמה מופחתת של תחרותיות, או בכך שבהם נמוך יחסית שיעור הנשירה, שמדרך הטבע מפחיתה את המתאם הנצפה. מוצע לשקול אומדן התקפות באוכלוסיה רק על בסיס מקדמי מתאם ממסלולים בהם המהימנות והסלקטיביות גבוהות. האומדן לתקפות ציון הסכם על-פי שיטה זו, על בסיס המדגם המנותח כאן, הוא 0.645.

מבוא

במרכז הארצי לבחינות ולהערכה נערכות באופן תקופתי בדיקות של תקפות החזאים המשמשים לברירת תלמידים לאוניברסיטאות כנגד מדדי ההצלחה בלימודים לאחר הקבלה (ראה, למשל, קנת-כהן, ברונר ואורן, 1998, 1999; Kennet-Cohen, Bronner & Oren, 1999). חישוב התקפות נעשה באמצעות מיצוע מקדמי מתאם במסלולי הלימוד השונים לאחר תיקונים לקיצוץ תחום בברירה ישירה ע"י נוסחת תיקון המשתמשת במנת סטיות התקן, בקבוצת המתקבלים ובאוכלוסיית המועמדים, של החזאי על-פיו מתבצע המיון (ראה, למשל, Guilford, 1965/1942; Linn, 1983; Lord & Novick, 1968).

במחקר מקדים (נבון וכהן, 2005) ניסינו לעיין מחדש בשאלת אומדנו המדויק של מקדם התקפות לאור ניתוח קפדני של חוסנו של התיקון הסטטיסטי המקובל בתחום. הניתוח העלה שכאשר לא מתקיימות ההנחות העומדות בבסיס התיקון, הוא עלול להניב אומד מוטא, בדרך-כלל כלפי מטה. זה מתיישב גם עם מסקנות מניתוחים מתמטיים המופיעים בכמה מאמרים בספרות הסטטיסטית (Greener & Osborn, 1980; Gross, 1982; Gross & Fleischman, 1987; Holmes, 1990; see review in Sacket & Yang, 2000). מעבודה שנעשתה לבקשתנו בהמשך לכך ע"י תלמידה לתואר שני בחוג לסטטיסטיקה באוניברסיטה העברית על נתונים שסופקו ע"י המרכז (וקולנקו-לגון, 2006), עולה (בהדגמה על חוג עתיר-תלמידים מסוים) שבניתוח חלופי שאינו מניח לינאריות והומוסקדסטיות כמובנות מאליהן מתקבל מקדם תקפות גדול ב- 17% מזה המתקבל בתיקון המקובל.

אך עיקרו של דוח זה נסב על בעיה בסיסית יותר באומדן התקפות מעבר לחוסן ההנחות הסטטיסטיות. הבה נניח שכל ההנחות הללו מתקיימות או לחילופין שהתיקון חסון לרמת הפרתן בפועל. האם במקרה זה סביר לאמוד את התקפות באופן המקובל?

בניתוח המדווח להלן מצאנו אישוש אמפירי לאפשרות שמקדמי התקפות הנאמדים בדרך המקובלת (0.55 עבור החזאי הנקרא ציון סכס, קרי צירוף ממוצע הבגרות וציון הבחינה הפסיכומטרית במשקל שווה; ראה קנת-כהן, ברונר ואורן, 1999) הם אומדני חסר.

הניתוח החל מתהייה אודות ההצדקה לאמידת התקפות באמצעות מיצוע פשוט מעבר לכל מסלולי הלימוד. תקפות, במובנה הנומינלי, מתייחסת ליכולתו של מבחן למדוד תכונה או כושר. במקרה הנדון, תקפות חזאי לפיו ממיינים מועמדים ללימודים אקדמיים היא יכולת החזאי למדוד

את הפוטנציאל להצלחה בלימודים. כמובן שהמדד התפעולי לתקפות מוגבל במידה שבה הקריטריון, כלומר, הצלחה בלימודים בפועל כפי שהיא נאמדת באמצעות ממוצע ציוני שנה א', מתקרב למיצוי אותו פוטנציאל. ההנחה התיאורטית היא שמתקיימים לפחות שלושה תנאים בקריטריון: א. הוא מבוסס על מדדים מהימנים ומייצגים המתואמים עם המבנה הנומינלי הנקרא "כושר אקדמי", ב. הציון אינו מושפע כמעט מכל מיני הטיות כמו אפליה מתקנת, רצון לרצות תלמידים וכו', ג. המוטיבציה של התלמידים להשיג ציונים גבוהים ככל האפשר היא לפחות כמו המוטיבציה להשיג ציון גבוה בחזאי.

ללא ספק, תנאים אלה אינם מתקיימים במלואם. כמו כן אין ספק, שקשה מאד לבדוק עד כמה הם מתקיימים בכל מסלולי הלימודים. מצד שני, ייתכן שניתן למצוא הבדל שיטתי בין סוגים מוגדרים של מסלולי לימודים.

ייתכן, למשל, שיש הבדל משמעותי במהימנות הקריטריון בין חוגים, למשל בין חוגים בעלי אופי שונה של הדיסציפלינה הנלמדת בהם. במידה שזמינות מדידות מהימנות (מה שאינו מתקיים בנתונים שעמדו לרשותנו) ניתן לבדוק זאת, ואף ניתן לתקן את המתאם עבור חוסר-מהימנות. מצד שני, המדידות המקובלות של מהימנות רגישות רק להיבט אחד שלה - עקביות ויציבות כפי שהן משתקפות במתאם בין מדדים (למשל בין ממוצע ציוני שנה א' לממוצע ציוני שנה ב'). אלא שמהימנותו של קריטריון במובן הרחב כוללת גם את שני התנאים האחרונים מבין השלושה שנמנו לעיל. ספק אם המדידות המקובלות משקפות את המידה בה אלה מתקיימים.

המידה בה מתקיימת מוטיבציה להישגים בלימודים (ראה תנאי ג' לעיל) תלויה מן הסתם פחות באופי הדיסציפלינה ויותר בסלקטיביות של מסלול הלימודים. אפשר שהלימודים בחוג סלקטיבי תחרותיים יותר ולכן מביאים יותר לידי ביטוי כישורים המתואמים עם החזאי אך רדומים בסביבה פחות תחרותית. אפשרות זו עולה מצירוף חמש ההנחות הבאות: (א) בחוגים לא-סלקטיביים רבים מתקיימת תועלת פוחתת והולכת משיפור ההישגים בלימודים, עד כדי יצירת אסימפטוטה בחלקו העליון של טווח ההישגים¹. (ב) כמו כן, ייתכן שלפעמים ניתן להגיע להישגים מרביים, או לפחות כאלה המניבים תועלת מרבית, באמצעות יכולת לא גבוהה במיוחד. (ג) אם א' נכונה, קל וחומר אם

1. בכל החוגים קיימים אמנם תמריצים, למשל מלגות לימודים. אך יש להניח שההנעה העיקרית של רוב התלמידים נובעת מגורמים פסיכולוגיים כגון עניין והישגיות, או מבוססת על סיכויי מעבר לשלבי לימוד מאוחרים יותר (שנה ב', תואר שני וכו').

גם ב' נכונה, אזי אם נניח שהישגים הם פונקציה מכפלתית של יכולת ומאמץ, המוכשרים ביותר יוכלו להשקיע פחות מאמץ ועדיין להגיע לתועלת מספקת. (ד) מאחר שהמוכשרים מודעים לכך, ולרבים מהם אין שום מניע אישי נלווה של תחרות בינם לבין עצמם, אלה אמנם ישקיעו את המאמץ המזערי הנחוץ להשיג הישג סביר בעיניהם, ועל כן יהיו לא מעטים שיגיעו לאותו הישג חרף הבדלים ביכולתם. (ה) משום ד', ייווצר קיצוץ תחום אלקטיבי בקצה העליון של התפלגות ההישגים, שיביא להפחתה במתאם הנצפה בין כל חזאי הצלחה לבין ההישגים עצמם, אך לא יקוּוּזוּ בתיקון המבוסס על הקיצוץ הנחזה בהתפלגות ציוני החזאי כתוצאה מהסלקציה.

במלים אחרות, כשהסף להצלחה סבירה בעיני רוב הלומדים ניתן להשגה בלי מיצוי הפוטנציאל, ההפרשים ביכולות הלומדים המסוגלים להשיג את הסף ישתקפו פחות ואולי יותר במידת המאמצים שיושקעו על ידם. במקרה זה ההבדל בהישגים בין תלמידים הנמצאים, למשל, במאון ה-90 ובמאון ה-80 על החזאי יהיה גדול הרבה יותר בסביבה תחרותית מאשר בסביבה לא תחרותית (שם אולי שניהם ישיגו ממוצע ציונים קרוב ל-100). המתאם בין יכולת להישגים יהיה אפוא קטן יותר בסביבה לא תחרותית².

מאידך, אפשרי שקבלה על פי סף גבוה מפחיתה את אחוזי הנשירה. אם כך הדבר, המדגמים על פיהם נאמדת התקפות סובלים יותר מנתונים חסרים בקצה התחתון במסלולים הלא סלקטיביים, ועל כן אומדן התקפות מהם מהווה אומדן-חסר.

ובכן, השונות בתקפות החזאי במסלולים שונים עשויה לנבוע משני סוגי סטיות מההנחה התיאורטית הרווחת לגבי הקריטריון. שתי האפשרויות מאפשרות להסביר שונות בממצאי תקפות בין מסלולים ללא התייחסות לחוסן נוסחת התיקון עצמה להפרת הנחותיה. ההסבר הראשון טוען שאומדן התקפות מנתוני מסלולים מסוימים עלול להיות אומדן חסר בשל מהימנות קריטריון נמוכה במיוחד. ההסבר השני טוען שאילו כל המועמדים היו מתחרים בסביבה תחרותית כמו זו שקיימת במסלולים הסלקטיביים, התקפות הייתה קרובה לזו המשתקפת במתאם המתקוקן לקיצוץ תחום עבור מסלולים אלה. שני ההסברים מעלים ספק לגבי דיוקם של הממצאים הקיימים אודות תקפות החזאים המשמשים למיון.

2. ההשוואה כאן היא בביצוע של אותם תלמידים בשתי סביבות שונות, תחרותית ופחות תחרותית. לא ניתן לגזור ממנה דבר על השוואה בין מידת ההשקעה ומובחנות ההישגים של אותם תלמידים לבין מידת ההשקעה ומובחנות ההישגים של תלמידים בעלי כשרים נמוכים יותר. באופן דומה, למרות שיש לשער שסביבה בלתי תחרותית גורמת להפחתת השונות, לא ניתן לבדוק זאת בהשוואה בין חוגים השזורה מן הסתם עם השפעות כושר ממוצע.

יחידת ניתוח במחקר זה היא קבוצת תלמידים שלמדו בחוג לימודים נתון במוסד מסוים ובמחזור נתון. דגמנו 778 יחידות ניתוח (הכוללות בס"ה 76,720 תלמידים) במסלולי לימוד שונים בשש האוניברסיטאות ובארבעה מחזורי לימוד (תשנ"ז, תשנ"ח, תשנ"ט, תשס"ג), וחישבנו עבור כל אחת מהן את המתאם הנצפה בין ציון הסכם המשמש לקבלה לאוניברסיטה (צירוף של ממוצע ציוני הבגרות והציון בבחינה הפסיכומטרית) לבין ממוצע הציונים בשנה א', את המתאם המתקן לקיצוץ תחום, וכן את המתאם הצפוי, שהוגדר כמתאם הנגזר מתפעול קיצוץ התחום שאירע בפועל על המתאם באוכלוסיה (באמצעות הנוסחה ההופכית לנוסחת התיקון), בהנחה שלאחרון ערך נתון (למשל, 0.6). לצורך התיקון לקיצוץ התחום השתמשנו באומדן סטיית התקן של ציון הסכם בקרב כלל המועמדים במחזורים תשנ"ב ותשנ"ג. מסלולי הלימוד מוינו לפי סלקטיביות (מוגדרת אופרציונלית על-פי סטיית התקן של החזאי בקבוצת הלומדים ביחידת הניתוח, כפי שהיא מופיעה בנוסחת התיקון לקיצוץ תחום, ברזולוציה של ארבע קטגוריות המובחנות על בסיס הרביעון בו נופלת אותה סטיית תקן בהתפלגות מעבר לכלל יחידות הניתוח³) ואופי הדיסציפלינה הנלמדת (מדעי הרוח והחברה לרבות משפטים, מדעים מדויקים⁴). אפקט של גורם הסלקטיביות ניתן לייחס

3. ליתר ביטחון נעשה ניתוח גם לפי הגדרה אופרציונלית אחרת - המאון ה-25 בהתפלגות ציוני הקבלה של התלמידים. התוצאות דומות למדי. בחרנו להציג את הניתוח לפי הגדרת הסלקטיביות כסטיית התקן בחזאי מאחר שהיא זו בה משתמשת נוסחת התיקון, ועל כן ניתן לגזור ממנה ניבויים חדים לגבי המתאמים המתוקנים (ראה להלן).

4. השתמשנו בקטגוריזציה זו, כי רק בה ניתן היה למצוא את כל רמות הסלקטיביות בכל קטגוריה של אופי דיסציפלינה. בפירוט על-פי חוגים (בשמותיהם הנהוגים במוסדות השונים): הקטגוריה הראשונה כללה את החוגים מקרא, אנגלית, ערבית, תיאטרון, אמנויות, בלשנות, צרפתית, היסטוריה, כללי רוח, פילוסופיה, לשון עברית, תרבות צרפת, לימודי א"י, מוסיקולוגיה, ספרות כללית, ספרות עברית, מחשבת ישראל, לימודי יהדות, אקדמיה למוסיקה, לימודי האיסלם, מדעי הרוח כללי, ספרות עם ישראל, קולנוע וטלביזיה, תולדות עם ישראל, המזה"ת בעת החדשה, לימודי המזרח התיכון, לימודים ספרדיים, היסטוריה של עם ישראל, תולדות התיאטרון, אמנות - יצירה, ארכיאולוגיה, בוגר כללי במדעי הרוח, היסטוריה של המזה"ת, היסטוריה של עם ישראל, לימודי ארץ-ישראל, לימודי מזרח אסיה, מוסיקה, צרפתית, שפה וספרות ערבית, תולדות האמנות, חינוך, כלכלה, חשבונאות, ניהול כלכלה, מינהל עסקים, חשבונאות וניהול, ניהול מלונאות ותיירות, כלכלה ומינהל עסקים, כלכלה חקלאית ומינהל בחקלאות, ניהול, יחסים בינלאומיים, משפטים, תקשורת, גיאוגרפיה, מדע המדינה, פסיכולוגיה, בוגר כללי במדעי החברה, סוציולוגיה, קומוניקציה, הוראת המדעים, משפטים משולב, קרימינולוגיה, עבודה סוציאלית, קוגניציה בוגר, היבטים התפתחותיים בחינוך, כלכלה-מדעי-המדינה-פילוסופיה, סוציולוגיה ואנתרופולוגיה, צירוף פסיכולוגיה ומדעי החיים, הוראה, הוראת טכנולוגיה ומדע, לימודי לשון ומגדר, לימודים כלליים, מדיניות ומינהל בחינוך, מדעי ההתנהגות, מערכות מידע וניהוליות, משולב מדעי החברה, פוליטיקה וממשל; הקטגוריה השנייה כללה את החוגים כימיה, רפואה, פיסיקה, מתמטיקה, רוקחות, אדריכלות, ביולוגיה, מדעי החי, סטטיסטיקה, גיאופיסיקה, מדעי המחשב, ארכיטקטורה, הגנת הצומח, הנדסת חשמל, הנדסת מזון, הנדסת תעשייה, הנדסת מכאנית, רפואת שיניים, הנדסה אזרחית, הנדסה חקלאית, הנדסת מתחבים, הנדסת מכונות, רפואה בסיסית, תזונה קלינית, הנדסה גיאודטית, לימודי הסביבה, תכנית משולבת במדעים מדויקים, מתמטיקה-פיסיקה, מדעי בעלי-חיים, מדעי כדור הארץ, גיאולוגיה ומינרלוגיה, מדעי-מוח בין-תחומי, הנדסת מחשבים ותכנה, מדעי הצומח בחקלאות, הנדסה אווירונאוטית, הנדסת תעשייה וניהול, ביוכימיה-תזונה, ביוכימיה מולקולרית, הנדסת מערכות תקשורת, הנדסה ביו-רפואית, הנדסה ביוטכנולוגיה, הנדסה גרעינית, הנדסה כימית, הנדסת בניין, הנדסת חומרים, הנדסת חשמל-פיזיקה, הנדסת מערכות מידע, הנדסת תוכנה, מדעי החיים, מדעי הקרקע והמים, מחשב וביוטכנולוגיה, תכנית משולבת מדעי החיים ורפואה, מתמטיקה ומדעי המחשב, ניהול משאבי מזון, שילוב הנדסה ומחשב.

לתחרותיות סביבת הלימודים. אפקט של גורם אופי הדיסציפלינה, כאשר הסלקטיביות מוחזקת קבועה, סביר לייחס בין היתר להבדלים במהימנות קריטריון.

במחקרים קודמים שנערכו במרכז הארצי לבחינות והערכה לצרכים אחרים (קנת-כהן, 2001; ברונר, 2004) התקבלו אפקטים על מידת תקפות הניבוי לשני גורמים אלה: נמצאה תקפות גדולה במידה ניכרת בחוגים בעלי סלקטיביות גבוהה מאשר בחוגים בעלי סלקטיביות נמוכה, וכן נמצאה תקפות גדולה במקצת בחוגים כמותיים (או ריאליים) מאשר בחוגים מילוליים (או הומניים). אך מאחר שאחת ממטרותינו העיקריות הייתה לבחון את אפקט הסלקטיביות באופן מדוקדק, השתמשנו במדד אופרציונלי שונה (על-פי סטיות תקן, בעוד שהמחקרים המצוטטים השתמשו במדד המוגדר על-פי הממוצע) ובעל רזולוציה גבוהה יותר (ארבע רמות, בעוד שהמחקרים המצוטטים השתמשו בשתיים). החשיבות של הגדרת הסלקטיביות באופן בו הגדרנוה היא בכך, שזו הגדרתה בנוסחה המקובלת לתיקון לקיצוץ תחום. אם ניתן היה לתהות שמא אפקט הסלקטיביות במחקרים קודמים על מקדמי המתאם המתוקנים אינו אלא ארטיפקט של תיקון-יתר או תיקון-חסר בחלק מהחוגים, תהייה זו בטלה כאשר הן הסלקטיביות והן התיקון עברה מתבססים על אותה הגדרה בדיוק. כמו כן, השתמשנו בדיכוטומיה שונה להגדרת משתנה אופי הדיסציפלינה (למשל, בעוד כלכלה ומתמטיקה נפלו במחקרים המצוטטים באותה קטגוריה, במחקר הנוכחי הם סווגו לקטגוריות שונות).

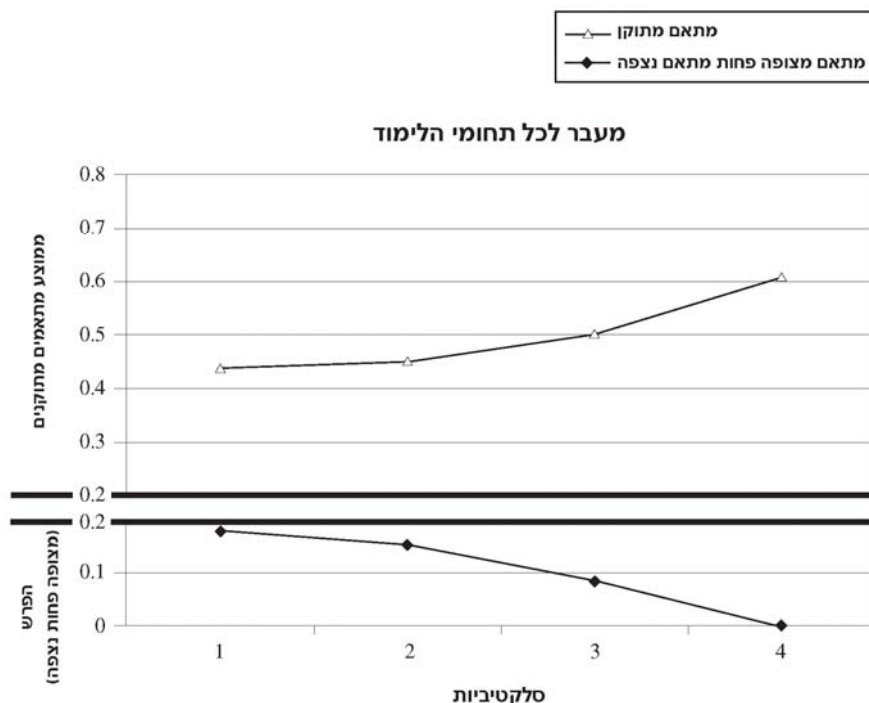
תוצאות ודיון

מהשערת האפס שהמתאם הנצפה במסלול מושפע אך ורק ממקדם המתאם האחיד באוכלוסיה ומקיצוץ התחום באותו מסלול נובעים שני ניבויים לגבי השפעת הסלקטיביות: א. המתאם הנצפה יפחת ככל שהמסלול סלקטיבי יותר באותו שיעור כמו המתאם הצפוי, ב. המתאם המתוקן לא ישתנה כלל כפונקציה של סלקטיביות.

בציור 1 מוצגים עבור ארבע רמות הסלקטיביות (כאשר הסלקטיביות הגבוהה ביותר מצוינת בספרה 4), מעבר לכל המסלולים, גרפים של ממוצעי מתאמים⁵ מתוקנים (בחלקו העליון של הציור), וכן של הפרשים בין ממוצעי מתאמים מצופים (כשהמתאם באוכלוסיה הוא 0.6) ונצפים (בחלקו התחתון של הציור).

5. הממוצעים, כאן ולהלן, חושבו על מקדמים שהותמרו בעזרת נוסחת פישר-Z (ראה, למשל, Guilford, 1965).

ציור 1. ממוצעי מתאמים מתוקנים (בגרף העליון) והפרשים בין ממוצעי מתאמים מצופים (בהנחה שהמתאם באוכלוסייה הוא 0.6) ונצפים (בגרף התחתון) כפונקציה של רמות סלקטיביות (מעבר לכל המסלולים).



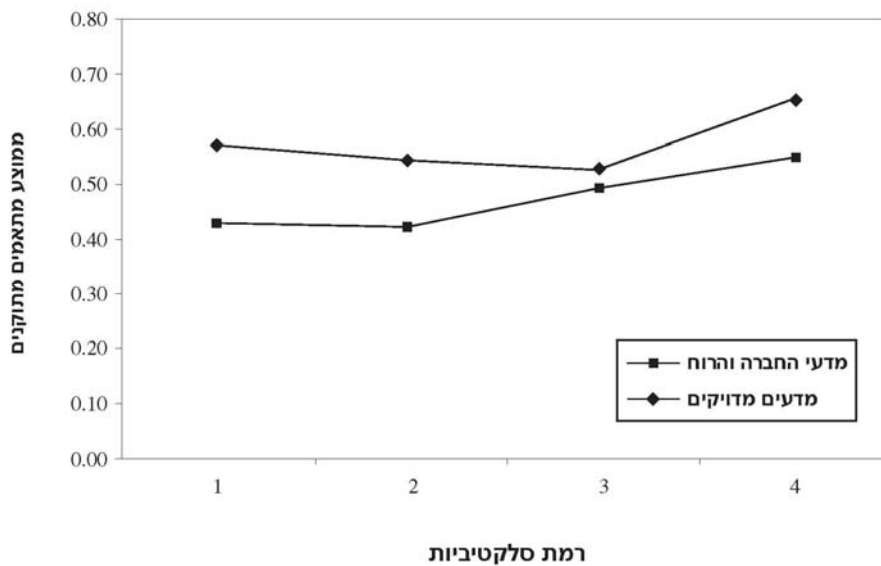
ממוצעי המתאמים המתוקנים בכל יחידת ניתוח נותחו בניתוח שונות בהליך GLM, מאחר שמספרי הנתונים בתאים לא היו שווים. נמצא כי בניגוד לניבוי ב', המתאם המתוקן גבוה יותר במסלולים סלקטיביים. האפקט הראשי של גורם הסלקטיביות נמצא מובהק, $F(3,770)=10.51, p<.0001$. בניתוחי פוסט-הוק בעזרת מבחן דאנקן נמצא שרמת הסלקטיביות הגבוהה נבדלת משלוש הרמות האחרות באופן מובהק, אך אין הבדלים מובהקים בין יתר הרמות לבין עצמן. אפקט זה מתיישב עם ההשערה שמיצוי פוטנציאל התקפות מתאפשר פחות כשפוחתת הסלקטיביות בשל תחרותיות נמוכה.

הגרף המתאר את הפרשים בין ממוצעי המתאמים המצופים והנצפים ממחיש שמקור השונות בין רמות סלקטיביות במתאם המתוקן הוא בכך, שככל שהמסלול פחות סלקטיבי המתאם הנצפה נופל יותר מהצפוי (בניגוד לניבוי א'). עצם העובדה שיש הפרש כזה מצביעה ככל הנראה על בעיית קריטריון. מהממצאים עולה שבעיה זו חריפה יותר ככל שהסלקטיביות פוחתת. אם אמנם מקדם המתאם באוכלוסייה הוא 0.6, אזי הבעיה אינה קיימת כלל ברמת הסלקטיביות הגבוהה, אך

מופיעה בשלוש הרמות הנמוכות יותר, במידה התלויה מונוטונית בסלקטיביות. תופעה זו ניתנת להסבר ע"י הפחתת המוטיבציה בשל תחרותיות נמוכה.

היבט אחר של הנתונים הוא המידה בה הם מושפעים מאופי הדיסציפלינה. מאחר שגורם זה אורתוגונלי במערך המחקר לגורם הסלקטיביות, אפקט ראשי שלו עשוי לנבוע מהבדלים במהימנות הקריטריון. כמו כן נבחן גורם זה כדי לבדוק אם תקפותם של ממצאים לגבי השפעת הסלקטיביות אינה תלויה באופי הלימודים, לפחות כפי שמשקף בהבדל בין לימודי המדעים המדויקים ללימודי כל יתר הדיסציפלינות. בציר 2 מוצגים גרפים של ממוצעי מתאמים מתוקנים עבור 8 צירופים של אופי דיסציפלינה ורמת סלקטיביות. בלוח 1 מוצגים בנוסף לממוצעים גם מספרי התצפיות (יחידות הניתוח) עליהן הם מבוססים.

ציר 2. ממוצעי מתאמים מתוקנים כפונקציה של רמות סלקטיביות (באבסציסה) ואופי הדיסציפלינה (בפאנל).



לוח 1. ממוצעי מתאמים מתוקנים בין ציון הסכם לקריטריון (מספרי תצפיות בסוגריים) בכל אחד מהצירופים של רמת סלקטיביות ואופי דיסציפלינה.

אופי דיסציפלינה

רמת סלקטיביות	מדעים מדויקים	מדעי הרוח והחברה
1 נמוכה	0.555 (19)	0.432 (176)
2	0.513 (53)	0.421 (141)
3	0.505 (98)	0.488 (97)
4 גבוהה	0.645 (120)	0.567 (74)

בניתוח השונות על ממוצעי המתאמים המתוקנים נמצא שמלבד האפקט המובהק של גורם הסלקטיביות, יש השפעה מובהקת לגורם אופי הדיסציפלינה, $F(1,770)=25.97, p<.0001$, אך האינטראקציה בין שני הגורמים אינה מובהקת, $F(3,770)<1$. בניתוח עוקב, נמצאו אפקטים פשוטים לגורם הסלקטיביות גם בתוך כל רמה של אופי דיסציפלינה: $F(3,484)=7.80, p<.0001$ במדעי הרוח והחברה, $F(3,286)=4.44, p<.005$ במדעים המדויקים. המתאם המתוקן גבוה יותר במסלולים הסלקטיביים במיוחד (אם כי אין ביטחון שהאפקט לגבי המסלולים האחרים מונוטוני בסלקטיביות⁶).

כאמור, את האפקט הראשי של אופי הדיסציפלינה ניתן לייחס, לפחות בחלקו, להבדלים במהימנות המדדים המשמשים לחישוב הקריטריון. ציוני מבחנים בהם אין כמעט מחלוקת בין מצייננים שונים על נכונותן של התשובות, משום קיומם של פתרונות מוסכמים, צפויים להיות

6. לכאורה ניכרת אף חריגה מסוימת ממונוטוניות חזקה, אך זו איננה מובהקת. החריגה הבולטת לעין היא במסלולים הפחות סלקטיביים במדעים המדויקים, שמספרם קטן יחסית (19 לעומת 53, 98 ו-120 ברמות הסלקטיביות הגבוהות יותר בהתאמה). כפי שניתן לראות מצויר 1, כאשר ממוצעי המתאמים ברמות הסלקטיביות השונות מבוססים על מספר כמעט שווה (± 1) של תצפיות, אין כלל חריגה כזו.

מהימנים יותר מציוני מבחנים בהם הערכתן של התשובות תלויה במידה ניכרת גם בעמדות או דרישות שאינן משותפות למציינים ובהתרשמות סובייקטיבית. סביר מאד להניח ששיעור המבחנים מהסוג הראשון נמוך יותר במדעי הרוח והחברה מאשר במדעים המדויקים.

השלכות

הממצאים מספקים חיזוק לטענה שלא כל מתאם נצפה במסלול לימוד כלשהו הוא בעל תרומה שווה למדידת התקפות במובנה המעניין, קרי עד כמה מסוגל החזאי לנבא הישגים אקדמיים בתנאים בהם מדידתם מהימנה (לפחות כמדידת החזאי) ולכל התלמידים יכולת מספקת לעמוד בדרישות הלימודים והנעה מספקת להישג מרבי (לפחות כמדידת הנעתם להשיג ציון גבוה בחזאי). המתאם המתוקן גבוה במובהק במסלולים בהם קיימת תחרותיות גבוהה וניתן להניח שמהימנות הקריטריון גבוהה מאשר בכל יתר המסלולים. עובדה זו לכשעצמה מצביעה על כך, שזה פחות או יותר האומדן הנכון (0.645 במדגם זה) לפוטנציאל התקפות על פי הגדרתה הנומינלית. המתאמים הנמוכים יותר במסלולים אחרים נובעים מן הסתם מבעיה כלשהי בטיב הקריטריון, אם בשל העדר תחרותיות ואם בשל מהימנות נמוכה. הסבר אפשרי אחר, שהחזאי מוטה לרעת תכנים הנלמדים במדעי הרוח והחברה, אינו נראה סביר לאור הגיוון הרב במרכיבי ציון הבגרות וההרכב המאוזן של הבחינה הפסיכומטרית (אנגלית, מתמטיקה וחשיבה מילולית). מיצוע המתאמים מכל המסלולים, לרבות אלה בהם איכות הקריטריון פחותה, איננו תורם אפוא לשיפור המדידה של תקפות החזאי. לדוגמה, האומדן המתקבל ממיצוע כל המסלולים במדגם זה⁷ הוא, הנמוך ב-0.141 (שהם כ-22%) מהאומדן על-פי המסלולים בהם איכות הקריטריון היא הגבוהה ביותר. מאחר שמדגם זה נדגם לצורך המחקר בלבד, האומדן המתקבל בו איננו תקף בהכרח לאוכלוסיית המועמדים כולה. כדי למצוא את האומדן הטוב ביותר מוטב לכלול בניתוח את כל מסלולי הלימוד (מכל המחזורים וכל האוניברסיטאות) העונים לדרישה הכפולה (תחרותיות גבוהה ומהימנות גבוהה).

7. אומדן זה נעשה ללא תיקון לחוסר מהימנות הקריטריון. אילו נעשה התיקון, כשם שנעשה ע"י קנת-כהן, ברונר ואורן (1999), סביר שהתוצאה הייתה במקצת גבוהה יותר.

מראי מקום

- ברונר, ש. (2004). השלכות שילוב מכפ"ל או מצרף בנערכת המיון להשכלה הגבוהה על תוקף הניבוי וההוגנות של המערכת (דוח טכני מס' 152). ירושלים: מרכז ארצי לבחינות ולהערכה.
- וקולנקו-לגון, ב. (2006). בדיקה מחדש של תקפות המיון לאוניברסיטאות לאור חוסנו של התיקון לקיצוץ תחום. עבודת מחקר, המעבדה לייעוץ סטטיסטי, האוניברסיטה העברית, ירושלים.
- נבון, ד. וכהן, י. (2005). בדיקה מחדש של תקפות המיון לאוניברסיטאות לאור חוסנו של התיקון לקיצוץ תחום (נייר עבודה מס' 1). ירושלים: מרכז ארצי לבחינות ולהערכה.
- 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. (1998). תוקף צירופים שונים של מרכיבי המבחן הפסיכומטרי ותעודת הבגרות בניבוי הישגי שנה א' באוניברסיטאות בישראל (דו"ח מס' 249). ירושלים: מרכז ארצי לבחינות ולהערכה.
- קנת-כהן, ת., ברונר, ש., ואורן, כ. (1999). ניתוח-על של תוקף ניבוי של מרכיבי מערכת המיון לאוניברסיטאות בישראל כלפי מידת ההצלחה בלימודים. מגמות, מ, 54-71.
- Greener, J.M., & Osburn, H.G. (1980). Accuracy of corrections for restriction in range due to explicit selection in heteroscedastic and nonlinear distributions. Educational and Psychological Measurement, 40, 33-346.
- Gross, A.L. (1982). Relaxing the assumptions underlying corrections for restriction of range. Educational and Psychological Measurement, 42, 795-801.
- Gross, A.L., & Fleischman, L.E. (1987). The correction for restriction of range and nonlinear regressions: An analytic study. American Psychological Measurement, 11, 211-217.
- Guilford, J.P. (1965). Fundamental statistics in psychology and education (4th ed.). New York: McGraw-Hill (Originally published in 1942).
- Holmes, D.J. (1990). The robustness of the usual correction for restriction in range due to explicit Selection. Psychometrica, 55, 19-32.

- Kennet-Cohen, T., Bronner, S., & Oren, C. (1999). The predictive validity of the components of the process of selection of candidates for higher education in Israel (Report no. 264). Jerusalem: National Institute for Testing and Evaluation.
- Linn, R.L. (1983). Pearson selection formulas: implications for studies of predictive bias and estimates of educational effects in selected samples. Journal of Educational Measurement, 20, 1-15.
- Lord, F.M., & Novick, M.R. (1968). Statistical theory of mental test scores. Reading, MA: Addison-Wesley.
- Sackett, P.R., & Yang, H. (2000). Correction for range restriction: An expanded typology. Journal of Applied Psychology, 85, 112-118.