



מהימנות הבחינה הפסיכומטרית ותוקפה בניבוי הצלחה בלימודים אקדמיים

תמר קנת-כהן

מדור מחקר

ינואר 2016

ברבים מחוגי הלימוד באוניברסיטאות, מספר המועמדים שיכולים להיקלט קטן ממספר המבקשים ללמוד. מכאן עולה הצורך למיין את המועמדים ולבחור מהם את המתאימים ביותר. שאלת איכותם של הכלים המשמשים למיין המועמדים היא סוגיה חשובה בעיני המוסדות האקדמיים, בעיני המועמדים עצמם ובעיני החברה כולה. שניים ממדדי היסוד שלפיהם נמדדת איכותם של כלי מיין הם מהימנות ותוקף: **מהימנות** של כלי מדידה בודקת באיזו מידה מדידות חוזרות של אותו נבחן בעזרת אותו כלי יתנו תוצאות דומות זו לזו, ו**תוקף** עוסק בשאלה באיזו מידה הבחינה מודדת את מה שהיא מיועדת למדוד. המוסדות האקדמיים מעוניינים לקבל את המועמדים בעלי הסיכויים הגבוהים ביותר להצליח בלימודים, ובהתאם לכך, הציונים שכלי המיין מספק אמורים לשקף את סיכויי המועמדים להצליח בלימודים. לפיכך, קשר חיובי בין הציונים שכלי המיין מספק ובין ההישגים בלימודים הוא עדות לתוקף כלי המיין בניבוי ההישגים בלימודים. מסמך זה מספק נתונים תמציתיים על מהימנות הבחינה הפסיכומטרית ועל תוקפה בניבוי הצלחה בלימודים אקדמיים.

מהימנות

אחת הדרכים למדוד מהימנות מבוססת על בדיקת עקיבות פנימית, המבררת באיזו מידה השאלות במבחן מודדות אותו משתנה (או אותם משתנים). יש מדדים שונים לבדיקת העקיבות הפנימית של מבחן. הנתונים שיוצגו להלן חושבו בשיטת קודר-ריצ'רדסון, שיטה הזוהה לחישוב מקדם אלפא ומתאימה למבחנים המורכבים מפריטים שבהם הציון הוא דיכוטומי, כמו "נכון" או "לא נכון". ערכו של מקדם המהימנות נע בין 0 ל-1.00: מקדם של 1.00 מעיד על מבחן בעל מהימנות מושלמת, ומקדם של 0 מעיד על העדר מהימנות. כלל אצבע החל על מבחנים עתירי סיכון, כדוגמת הבחינה הפסיכומטרית, הוא שמהימנות המבחן צריכה להיות לכל הפחות 0.80.

בלוח 1 מוצגים מקדמי המהימנות של ציוני הבחינה הפסיכומטרית. נתוני המהימנות מתבססים על למעלה מ-50 מנוסחי הבחינה במתכונת החדשה, הכוללת מטלת כתיבה. בלוח מוצגים ערכי החציון של מקדמי המהימנות של כל נוסחי הבחינה שהועברו במועדים אלו.

לוח 1: מקדמי המהימנות של ציוני הבחינה הפסיכומטרית

ציונים כלליים*			ציונים בתחומי הבחינה		
כללי בדגש כמותי	כללי בדגש מילולי	כללי רב-תחומי	אנגלית	חשיבה כמותית	חשיבה מילולית
0.94	0.94	0.95	0.92	0.87	0.88

* בציון הרב-תחומי משקל הציון בתחום החשיבה המילולית ומשקל הציון בתחום החשיבה הכמותית כפולים ממשקלו של הציון באנגלית. בציון כללי בדגש מילולי משקלו של הציון בתחום החשיבה המילולית הוא פי שלושה ממשקל כל אחד מן הציונים האחרים. בציון כללי בדגש כמותי משקלו של הציון בתחום החשיבה הכמותית הוא פי שלושה ממשקל כל אחד מן הציונים האחרים.

הנתונים מראים שהבחינה הפסיכומטרית מהימנה מאוד.

תוקף ניבוי

האוניברסיטאות בישראל והמרכז הארצי לבחינות ולהערכה (מאל"ו), הפועל מטעמן, עורכים מחקרים שוטפים כדי לבדוק את יעילות הבחינה הפסיכומטרית ככלי לניבוי מידת ההצלחה בלימודים אקדמיים. במחקרים אלו נבדק הקשר בין ציוני הבחינה הפסיכומטרית, ממוצע ציוני הבגרות וציון הסכם (שהוא צירוף של ממוצע הבגרות והציון הפסיכומטרי במשקלות שווים) לבין הישגי הסטודנטים בלימודים אקדמיים, בתום שנה א' ובתום לימודי הבוגר.

המדד המקובל ביותר בבדיקת תוקף ניבוי של כלי מיון הוא מקדם המתאם של פירסון בין ההישג בכלי המיון לבין ההישג בלימודים, והוא מכונה גם מקדם התוקף של כלי המיון. ערכו של מקדם המתאם נע בין -1.00 ל- +1.00. מקדם מתאם של +1.00 מעיד על קשר חיובי מושלם, ומקדם מתאם של 0 מעיד על העדר קשר. כלל אצבע מקובל לפירוש מקדמי מתאם (בערכים מוחלטים) הוא שמתאם נמוך מ-0.10 נחשב זניח, מתאם בטווח של 0.10-0.29 נחשב נמוך, מתאם בטווח של 0.30-0.49 נחשב בינוני, ומתאם של 0.50 ומעלה נחשב גבוה (Cohen, 1988). חשוב לציין שמהימנות גבוהה היא תנאי לתוקף ניבוי גבוה.

אמידת מקדמי התוקף של כלי מיון מחייבת התמודדות עם הבעיה של קיצוץ תחום. אנו מתעניינים בתוקף הניבוי של כלי המיון בקרב כל המועמדים, אך מסוגלים לחשוב רק בקרב אותם מועמדים שהתקבלו והתחילו ללמוד, כיוון שרק לגביהם יש מידע על הישגים בלימודים. הבעיה היא שבקרב קבוצת הלומדים, טווח הציונים – הן בכלי המיון והן בהישגים בלימודים – הוא טווח מצומצם יותר, או "מקוצץ", מה שמקטין את מקדם התוקף הנצפה. כדי לאמוד את מקדם התוקף האמתי, המשקף את התוקף בקרב כל המועמדים, יש צורך בתיקונים סטטיסטיים המסתמכים על הקשרים הנצפים בין המשתנים בקרב הלומדים ועל נתוני התפלגות המשתנים הידועים בקרב כל המועמדים. בנתונים שיוצגו נעשה שימוש בנוסחת תיקון לקיצוץ תחום שמתאימה לאופי הנתונים (Gulliksen, 1987, chap. 12).

בלוח 2 מוצגים מקדמי התוקף¹ של כלי המיון בניבוי שני קריטריונים: ציון שנה א' לתואר בוגר (Oren, Kennet-Cohen, Turvall, & Allalouf, 2014) וציון הגמר לתואר בוגר (קלפר, טרוול ואורן, 2014). הממצאים בנוגע לציון שנה א' מבוססים על נתוני 100,863 סטודנטים בכל האוניברסיטאות בישראל ממחזורים תשס"ו (2005/6) עד תש"ע (2009/10). הממצאים בנוגע לציון הגמר מבוססים על נתוני 100,180 סטודנטים בכל האוניברסיטאות בישראל ממחזורים תשס"ב (2001/2) עד תשס"ז (2006/7).

לוח 2: מקדמי התוקף של כלי המיון בניבוי ציון שנה א' וציון הגמר לתואר בוגר

ציוני הבחינה הפסיכומטרית			כלי רב־תחומי	ממוצע הבגרות	סכם רב־תחומי	הקריטריון
ציונים בתחומי הבחינה						
אנגלית	חשיבה כמותית	חשיבה מילולית				
0.32	0.37	0.36	0.43	0.36	0.46	ציון שנה א' לתואר בוגר
0.31	0.33	0.34	0.41	0.38	0.47	ציון הגמר לתואר בוגר

הנתונים שלעיל מספקים עדות לתוקף הניבוי של הציון הכללי בבחינה הפסיכומטרית, ליתרוננו, במונחי תוקף ניבוי, על ממוצע הבגרות, ולתרומתו לניבוי הקריטריון מעבר לממוצע הבגרות – תרומה שמשתקפת בפער שבין תוקף הניבוי של ממוצע הבגרות לתוקף הניבוי של ציון הסכם. מקדמי התוקף בניבוי ציון שנה א', שמחושבים במאל"ו באופן שוטף, דומים למקדמי התוקף בניבוי ציון הגמר לתואר בוגר.

1 המתאמים שמוצגים במסמך זה חושבו בתוך יחידות עיבוד (חוג במוסד לימודים כלשהו ובמחזור כלשהו) ומוצעו באופן משוקלל (לפי מספר הסטודנטים ביחידת העיבוד).

בלוח 3 מוצגים מקדמי התוקף של כלי המיון בניבוי ציון שנה א' לתואר בוגר (Oren et al., 2014) לפי תחומי לימוד. בלוח מפורטים מקדמי התוקף של שלושת הציונים הכלליים בבחינה הפסיכומטרית (ציון כללי רב-תחומי, ציון כללי בדגש מילולי וציון כללי בדגש כמותי) ומקדמי התוקף של שלושת ציוני הסכם המחושבים על פיהם (סכם רב-תחומי, סכם בדגש מילולי וסכם בדגש כמותי).

לוח 3: מקדמי התוקף של כלי המיון בניבוי ציון שנה א' לתואר בוגר לפי תחומי לימוד*

ציוני הבחינה הפסיכומטרית						ממוצע הבגרות	ציוני סכם			תחומי לימוד
ציונים בתחומי הבחינה			ציונים כלליים				סכם בדגש כמותי	סכם בדגש מילולי	סכם רב-תחומי	
אנגלית	חשיבה כמותית	חשיבה מילולית	כללי בדגש כמותי	כללי בדגש מילולי	כללי רב-תחומי					
0.38	0.34	0.39	0.43	0.46	0.45	0.37	0.43	0.45	0.44	משפטים
0.34	0.34	0.39	0.41	0.43	0.43	0.38	0.47	0.49	0.48	רוח
0.33	0.33	0.37	0.40	0.42	0.42	0.27	0.40	0.41	0.41	חברה - מילולי
0.24	0.13	0.18	0.22	0.24	0.24	0.12	0.18	0.20	0.19	עזר - רפואה
0.24	0.31	0.24	0.35	0.30	0.33	0.30	0.36	0.34	0.35	רפואה
0.29	0.39	0.31	0.44	0.39	0.43	0.47	0.54	0.52	0.53	חברה - כמותי
0.38	0.50	0.42	0.54	0.48	0.52	0.45	0.58	0.55	0.57	טבע
0.31	0.42	0.37	0.47	0.43	0.45	0.46	0.54	0.52	0.53	הנדסה
0.32	0.37	0.36	0.43	0.41	0.43	0.36	0.46	0.46	0.46	כולם

*מודגש בצבע: התוקף הגבוה ביותר בשלישיית ציוני הסכם ובשלישיית הציונים הכלליים בבחינה הפסיכומטרית.

גם בחלוקה לתחומי לימוד מתקבלת עדות עקבית לתרומתו של הציון הכללי בבחינה הפסיכומטרית לניבוי הקריטריון מעבר לממוצע הבגרות: בכל תחומי הלימוד תוקף ציון הסכם גבוה מתוקף ממוצע הבגרות. נוסף על כך, מהנתונים עולה שלצורך ניבוי בתחומי הלימוד המילוליים יש יתרון קל לציון הכללי בדגש מילולי, ואילו לצורך ניבוי בתחומי הלימוד הכמותיים יש יתרון לציון הכללי בדגש כמותי.

לסיכום, הנתונים המוצגים לעיל מעידים על מהימנות ותוקף גבוהים (ביחס למערכות מיון בארץ ובעולם), והם תומכים בשימוש בבחינה הפסיכומטרית בתהליך הבררה ללימודים לתואר בוגר באוניברסיטאות בישראל. מידע מפורט יותר על המחקרים שממצאיהם מוצגים לעיל אפשר למצוא במקורות הרשומים להלן.

מקורות

קלפר, ד', טורוול, א' ואורן, כ' (2014). **תוקף הניבוי של כלי המיון לאוניברסיטאות בישראל מול ממוצע תואר בוגר** (דוח מס' 403). ירושלים: מרכז ארצי לבחינות ולהערכה.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Hillsdale, NJ: Erlbaum.

Gulliksen, H. (1987). *Theory of mental tests*. Hillsdale, NJ: Erlbaum. (Original work published 1950)

Oren, C., Kennet-Cohen, T., Turvall, E., & Allalouf, A. (2014). Demonstrating the validity of three general scores of PET in predicting higher education achievement in Israel. *Psicothema*, 26, 117-126.