

לתקנן או לא לתקנן?

על יתרונות שבשימוש בציוני תקן של פריטי שאלונים
ומבחנים בחישוב ציון כולל.

רז ליבוביץ

המכללה האקדמית תל-אביב יפו

פרופ' אייל גמליאל

המרכז האקדמי רופין

לירון צוברי

האוניברסיטה העברית
המרכז לטכנולוגיה חינוכית

מדידת תכונה

- תכונה לרוב נמדדת באמצעות מספר רכיבים (להלן, פריטים) שכל אחד מהם נבחר לייצג אותה ומהווה עדות לה.
- מציוני הפריטים מחושבת תוצאה מספרית יחידה (להלן, ציון כולל) שאמורה לשקף את התכונה הנמדדת.

הנחת האקוויוולנטיות הפונקציונאלית

ווקסלר (Wechler, 1958) התייחס לסוגיית חישוב הציון הכולל במבחן למדידת אינטליגנציה ממספר תת-מבחנים:

"הצירוף של מגוון של מבחנים למדד יחיד של אינטליגנציה (הציון הכולל)... מניח בהכרח אחידות פונקציונאלית או אקוויוולנטיות בניהם... האקוויוולנטיות הפונקציונאלית של פריטי המבחן, היא תנאי הכרחי לתיקוף האריתמטיקה שבה משתמשים כדי להגיע למדד יחיד של אינטליגנציה. אריתמטיקה זו כוללת הקצאה של ערך מספרי לכל תשובה של המשתתף, צירוף או מיצוע של כל המספרים לכדי ציון כולל, והתייחסות לציון הכולל כאל המדד של המשתנה." (קאהן, 1998).

הנחת האקוויוולנטיות הפונקציונאלית

- בהקשר זה אילו הפריטים, השאלות, המטלות או תת-המבחנים היו נתפסים כמייצגים ישויות שונות, לא ניתן היה לחבר את המספרים שהוקצו להם כדי לבטא משתנה/תכונה כללית.
- לפיכך, מבחינה תיאורטית, לדעת ווקסלר משקלם של הפריטים המרכיבים את המבחן צריך להיות שווה בחישוב הציון הכולל למען המטרה של מדידת תכונה כללית. (קאהן, 1998).
- תקנון הציונים הגולמיים בתת-המבחנים המרכיבים את מבחן האינטליגנציה של ווקסלר, ומיפויים לסולם אחיד עם ממוצע 10 וסטיית תקן 3 נועדו להשיג מטרה זו.

השיטה הנהוגה לחישוב הציון הכולל

- בפועל, הנוהג הוא לחשב ציון כולל באמצעות סכימה או מיצוע של ציוני הגלם של הפריטים (להלן, ציון כולל המבוסס על ציוני הגלם של הפריטים).

הסתירה בין השיטה הנהוגה לחישוב הציון הכולל להנחת האקוויוולנטיות של פריטים

- חישוב הציון הכולל בשיטה הנהוגה סותר את ההנחה העומדת בבסיס חישובו של הציון הכולל; זאת מכיוון שכאשר הוא מחושב מציוני הגלם של הפריטים, כל פריט מקבל משקל שונה בחישוב הציון הכולל, כפונקציה של הפיזור של ציוני הפריט.
- ככל שהפיזור של ציוני הפריט יהיה גבוה יותר, כך הפריט ישפיע יותר על הציון הכולל וינבא יותר משונות הציונים הכוללים בין המשתתפים.
- במילים אחרות, ככל שסטיית התקן של הפריט תהיה גבוהה יותר, כך הוא יקבל משקל גבוה יותר בחישוב הציון הכולל.

דוגמה 1

התפלגות תשובות המשתפים בשני פריטים המודדים את
משתנה Y

משתתף	פריט א	פריט ב	ציון כולל
1	7	9	16
2	7	8	15
3	7	7	14
4	7	6	13
5	7	5	12

חישוב הציון הכולל באופן התואם את הנחת האקוויוולנטיות של הפריטים

- כדי לנטרל את ההשפעה הדיפרנציאלית של סטיות התקן של הפריטים בחישוב הציון הכולל, ניתן להמיר את ציוני הגלם של הפריטים בציוני תקן, וכך להביא את ציוני הפריטים לכדי סולם משותף בו לכל הפריטים ממוצע השווה ל-0 וסטיית תקן השווה ל-1 (להלן, ציון כולל המבוסס על ציוני התקן של הפריטים).
- בדרך זו הפריטים משוקללים באופן שווה בחישוב הציון הכולל (Stevens & Aleamoni, 1986).

דוגמה 2

התפלגות תשובות המשתתפים בשני פריטים המודדים את
משתנה Y

ציוני תקן			ציוני גלם			משתתף
ציון כולל	פריט ב	פריט א	ציון כולל	פריט ב	פריט א	
1.11	1.26	0.96	9.00	9	9	1
0.79	0.63	0.96	8.50	8	9	2
-0.72	0.00	-1.43	7.00	7	7	3
-0.44	-0.63	-0.24	7.00	6	8	4
-0.75	-1.26	-0.24	6.50	5	8	5
0.00	0.00	0.00	7.60	7.00	8.20	ממוצע
0.89	1.00	1.00	1.08	1.58	0.84	סטיית תקן

המשך דוגמה 2

מטריצת המתאמים בין הפריטים לציון הכולל

ציוני תקן		ציוני גלם	
ציון כולל		ציון כולל	
0.89	פריט א	0.80	פריט א
0.89	פריט ב	0.95	פריט ב

שיטה

- 64 משתנים שנמדדו בשאלוני דיווח עצמי ב-21 מחקרים ("נמוכי סיכון")
- ארבעה מבחני הישגים של ארבעה קורסי חובה מתודולוגיים בתואר ראשון שהועברו במוסד להשכלה גבוהה בישראל ("עתירי סיכון")

נמוכי סיכון	עתירי סיכון	
2656	486	סה"כ משתתפים
ליקרט	דיכוטומי	סולם הפריטים

שיטה

- עבור כל משתנה חושבו:

– ציוני גלם של הפריטים המרכיבים את כלי המדידה:

$(Q_1, Q_2, Q_3 \dots Q_j)$

– ציוני תקן של הפריטים המרכיבים את הכלי המדידה:

$(Z_{Q_1}, Z_{Q_2}, Z_{Q_3} \dots Z_{Q_j})$

– ציון כולל המבוסס על ציוני הגלם של הפריטים:

$ME[X] / SUM[X]$

– ציון כולל המבוסס על ציוני התקן של הפריטים:

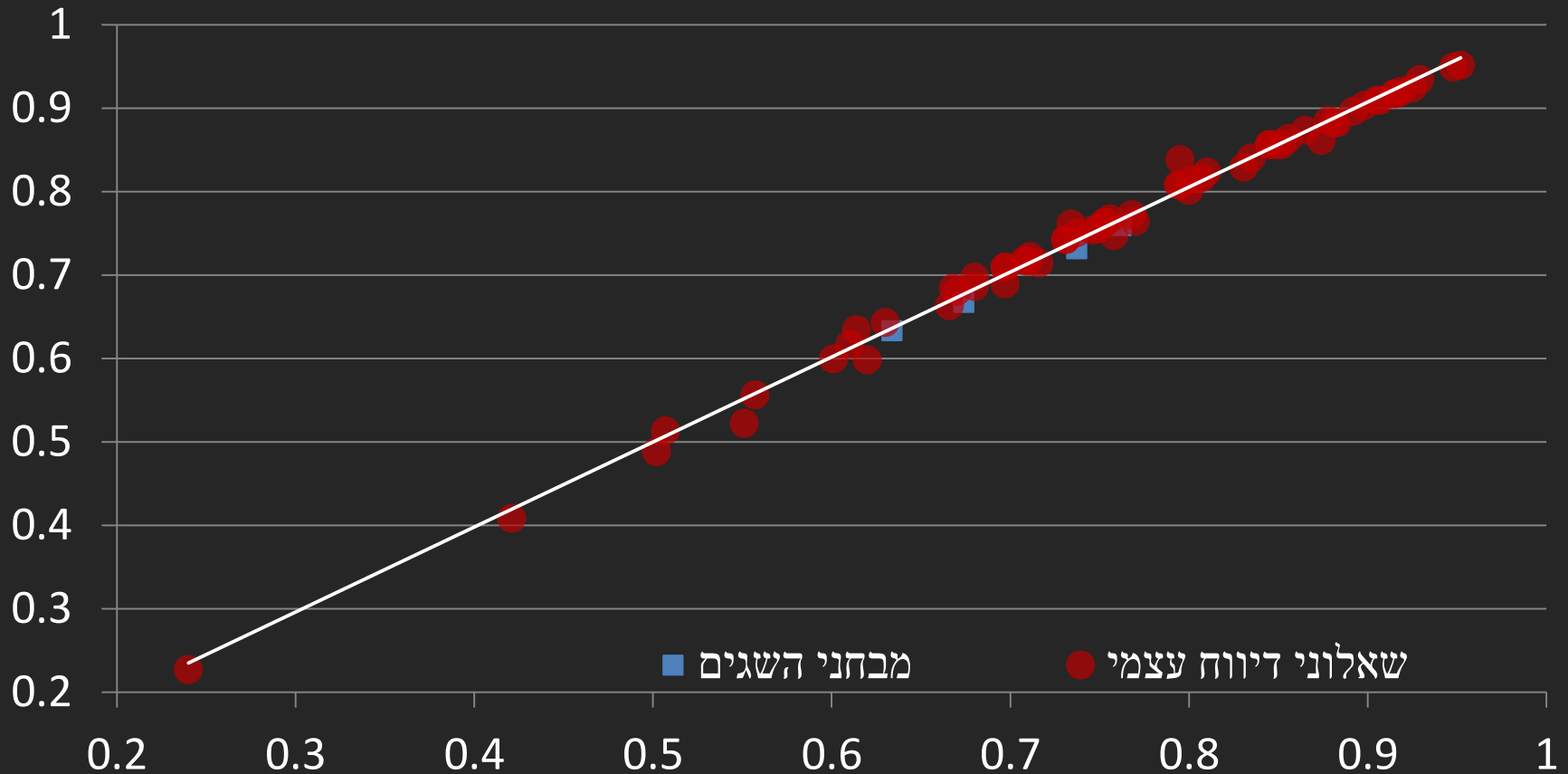
$ME[Z_x] / SUM[Z_x]$

שיטה

עתירי סיכון	נמוכי סיכון	
$(\alpha_x; \alpha_{zx})$		מהימנות
$(r(x,y) ; r(zx,y))$	$(r(x,y) ; r(zx,zy))$	תוקף
$R_x ; R_{zx}$ $\Delta = R_x - R_{zx} $ $ME[\Delta]/N$		הוגנות

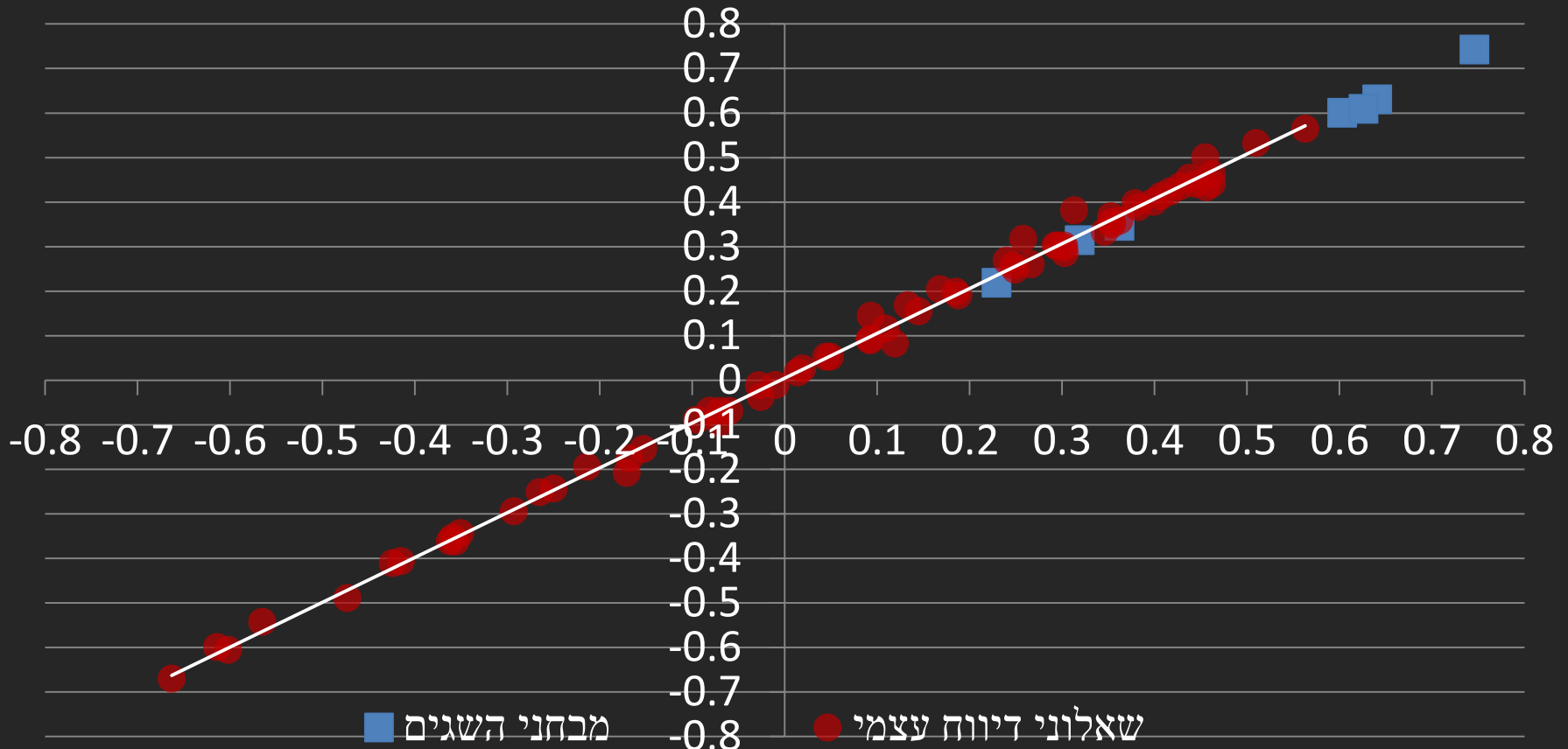
תוצאות - מהימנות

העקביות הפנימית של 64 שאלוני הדיווח העצמי ושל ארבעת מבחני ההישגים, המחושבת על סמך ציוני התקן של הפריטים (בציר Y) כפונקציה של העקביות הפנימית שלהם המחושבת על סמך ציוני הגלם של הפריטים (בציר X).



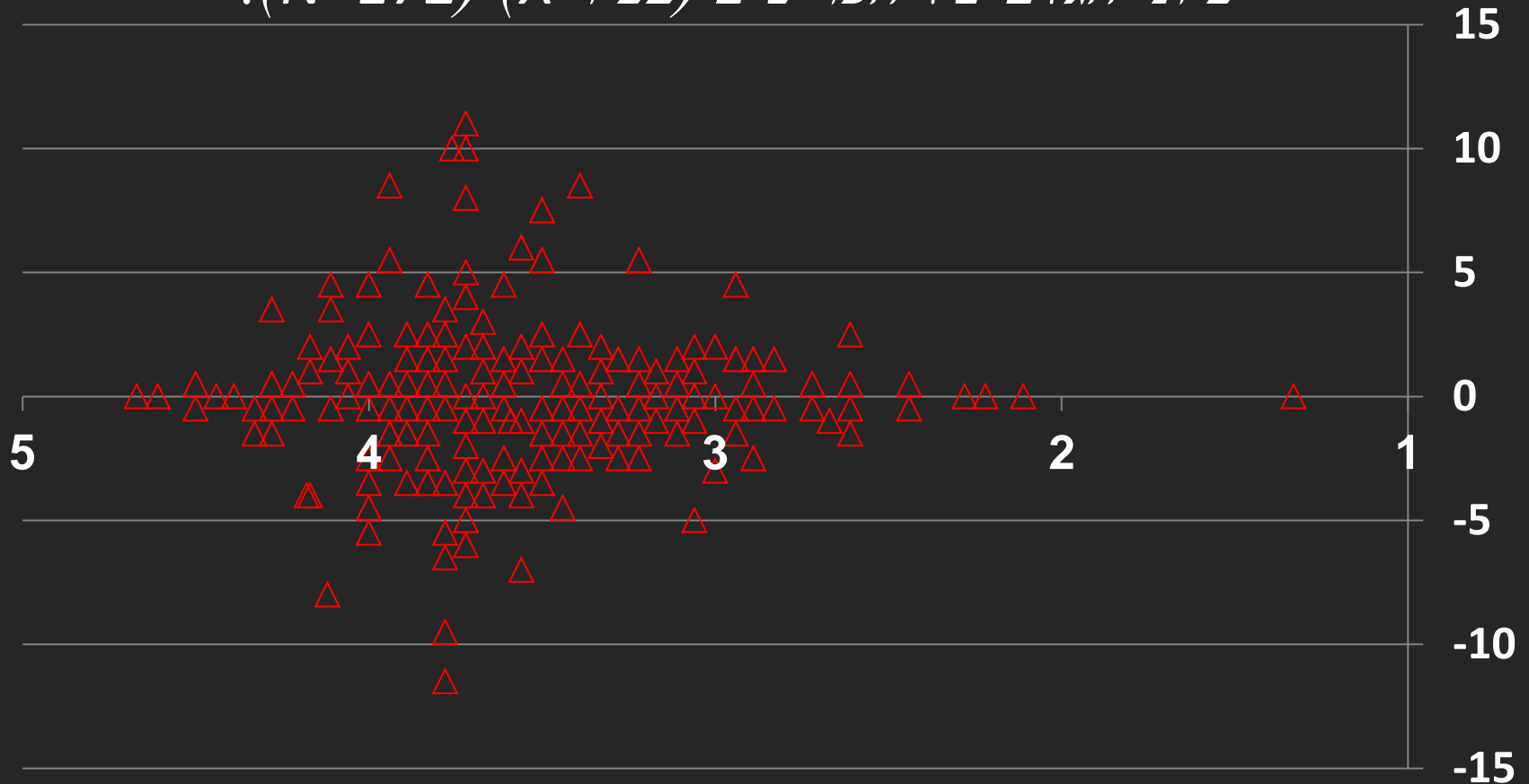
תוצאות - תוקף

מקדמי התוקף של שאלוני דיווח עצמי ומבחני ההישגים המחושבים על סמך הציונים הכוללים שחושבו מציוני הגלם של הפריטים (בציר Y) כפונקציה של מקדמי התוקף שלהם בהתבסס על ציוני התקן של הפריטים (בציר X).



תוצאות - הוגנות

ההפרש בין הדרוגים של שני הציונים הכוללים באחד משאלוני הדיווח העצמי (בציר Y) כפונקציה של הציון הכולל המחושב בהתבסס על ציוני הגלם של הפריטים (בציר X) ($N=172$).



דיון

- ממצאי המחקר אינם מצביעים על יתרון פסיכומטרי לאחת משתי שיטות הישוב הציון הכולל מבחינת שיקולי מהימנות או תוקף.
- השיטה הנהוגה לחישוב הציון הכולל "הוגנת" פחות משיטת החישוב המוצעת במובן שהציון הכולל המבוסס על ציוני הגלם של הפריטים מעניק משקל יתר לחלק מהפריטים ומשקל חסר לפריטים אחרים. זאת, לא באמצעות החלטה מושכלת אפרורית, אלא בדיעבד בעקבות פיזור שונה של הפריטים, ובניגוד לרציונל המובלע או הגלוי של מתן משקל שווה למרכיבים השונים.

דיון

- ממוצעים שמבוססים על ציונים כוללים (המחושבים מציוני גלם) המופקים ממדגמי משתתפים של מחקרים שונים או ממדגמים שונים באותו מחקר (כמו, מדגם נשים לעומת מדגם גברים) אינם באמת מקבילים, אלא במדגמים שונים פריטים של אותו שאלון מקבלים משקל שונה בחישוב הציון הכולל.

- חישוב ציון כולל המבוסס על ציוני התקן של הפריטים מבטיח מעבר למדידה מדויקת והוגנת יותר של משתנים גם את האפשרות להשוות בין תוצאות של מחקרים שונים שעושים שימוש בשאלונים מקבילים.

דיון

- שיטת החישוב המוצעת של הציון הכולל המבוסס על ציוני התקן של הפריטים ישימה בעיקר במבחנים תלויי-נורמה (norm-referenced test) לעומת מבחנים תלויי-קריטריון (criterion-referenced test).
- עם זאת לעיתים עושים שימוש יחסי בציונים כוללים המופקים ממבחנים, כמו למשל כאשר מקצים תלמידים לקבוצות לימוד לפי הישגים.

דיון

- לחישוב הציון הכולל באמצעות ציוני התקן של הפריטים או תת-המבחנים ישנה עדיפות ברורה כאשר מטרת המבחן היא להבחין בין נבחנים ולדרגם ביכולת או בתכונה כלשהי (כלומר כאשר מחושב ציון נורמה) על מנת למיינם, לסננם או לברור מתוכם.

תודה על ההקשבה