

מבחנים פסיכולוגיים

במצגת זו

- מבחנים פסיכולוגים – דוגמאות
- מה בודקים מבחנים פסיכולוגיים?
- מתי מבחן פסיכולוגי תקין ומתי אינו תקין?

מה משותף ל(כמעט) כל המבחנים

- מטלה המורכבת פריטים
- יש דרך לתרגם את הביצוע בפריט לציון
- הסכום (או הממוצע) של ציוני הפריטים מהווה את הציון במבחן

דוגמה 1: מבחן חרדה תכונתית

- המטלה: לאמר עד כמה היגד מסויים מבטא אותי בסולם 1-4
 - "אני מהיר חימה"
 - התשובה היא הציון בפריט
 - "אני אדם יציב"
 - התשובה **המהופכת** היא הציון בפריט
 - "אני מרגיש לא שווה"
 - התשובה היא הציון בפריט
- הציון במבחן הוא סכום הציונים בפריטים

דוגמה 2: מבחן אריתמטיקה

$$8+15 = ? \cdot$$

$$24:2 = ? \cdot$$

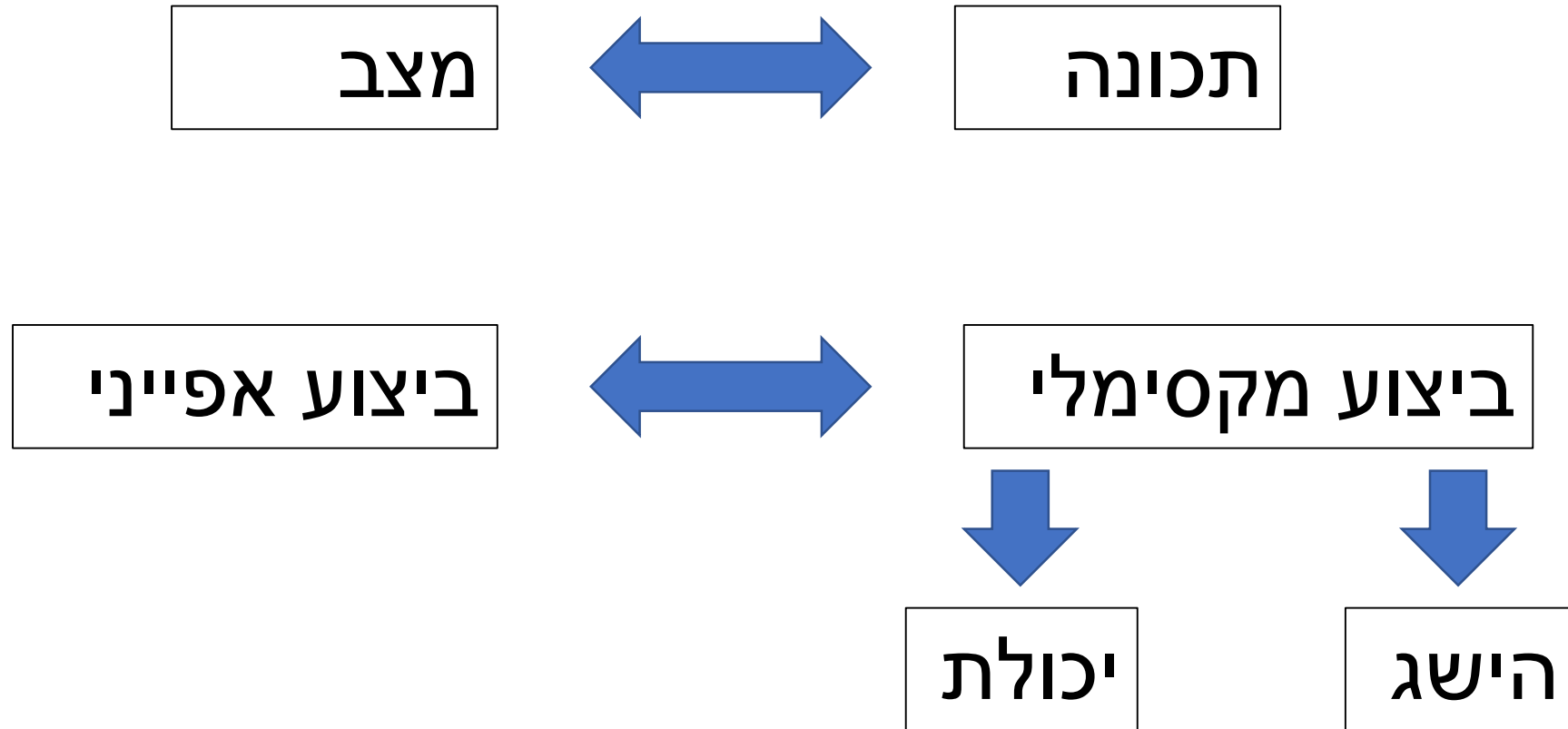
• הציון בפריט הוא 1 אם הפתרון נכון ו 0 אם אינו נכון

• הציון במבחן הוא סכום ציוני הפריטים

(כמעט) כל התופעות שנמדדות הן סטטיסטיות באופיין

- גם אם אתם שולטים במיומנות קוגניטיבית, ייתכן שתטעו לפעמים
- האמדן מבטא את הסיכוי להצליח
- גם למי שאינו סובל מדיכאון יש לפעמים חוויה דכאונית
- מה שהופך את המחשבות או החוויות לדיכאון הוא הסבירות להתרחשותן
- התערבויות (טיפול) יעילות כאשר הן מורידות את הסבירות לחוויות או התנהגויות לא רצויות.
- כדי שהמבחן יהיה טוב יש לוודא כי דגמנו מספיק התנהגויות (ואת ההתנהגויות הנכונות)

סוגי מבחנים פסיכולוגיים



מתי מבחן פסיכולוגי תקין ומתי אינו תקין?

• מהימנות לסוגיה

- האם הציון במבחן תלוי במי בדק את המבחן?
- האם המבחן כולל דגימה גדולה ומגוונת מספיק של התנהגויות כדי לשקף את מה שנמדד?
- האם המבחן בודק תכונה יציבה?

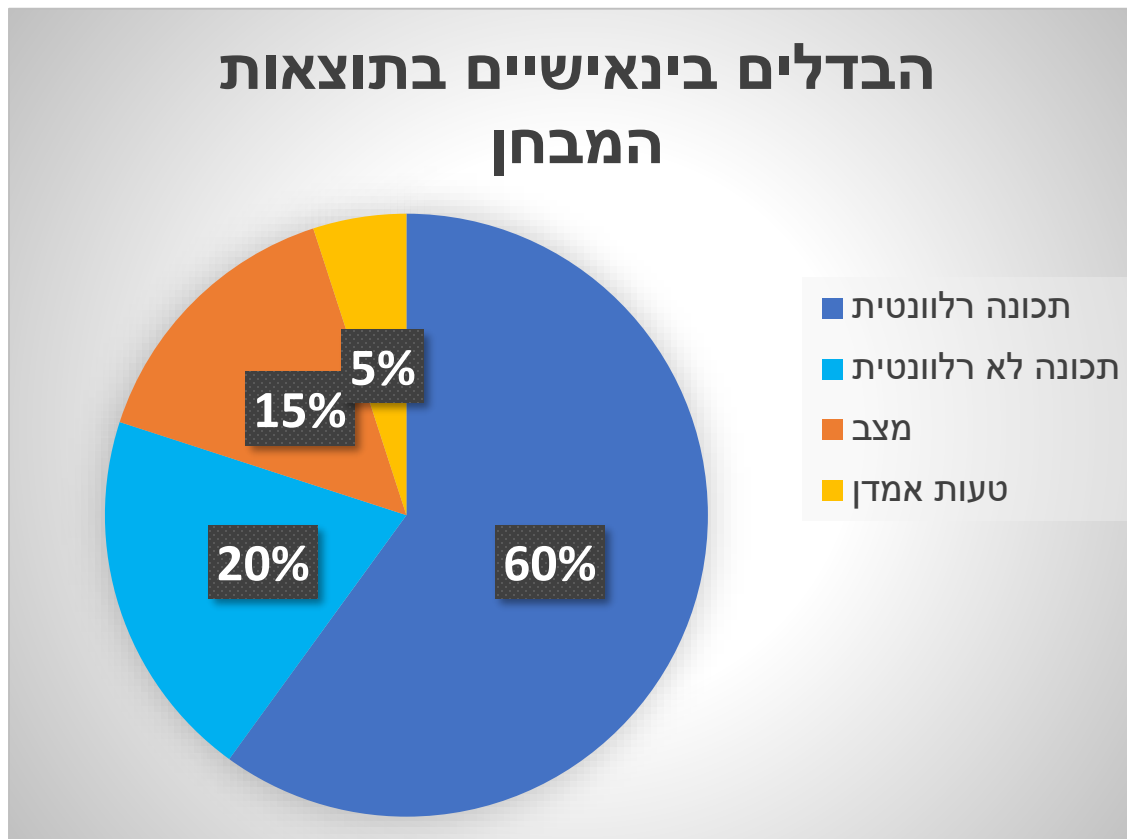
• תוקף לסוגיו

- האם הציון במבחן מנבא התנהגות עתידית רלוונטית?
- מה היא המשמעות של ציון גבוה או נמוך במבחן?
- **נורמות** - האם יש דרך לקבוע לגבי ציון אם הוא גבוה או נמוך יחסית?

מהימנות ביו שופטים: האם הציון תלוי בנותן הציון?

- נבדק באמצעות בדיקת מידת ההסכמה של שניים או יותר נותני-ציונים (שופטים)
- מידת ההסכמה נבדקת באמצעות מדדים כמו:
 - אחוז הסכמה
 - Intra-Class Correlation-ICC
- הבעיה לא קיימת (למעשה) עם משתנים שנבחנו בשאלות סגורות ("אמריקאיות"), למשל, או במדדים שבהם אין מימד של התרשמות או הערכה (זמן, גובה, משקל, וותק בעבודה, וכדומה)

גם כשיש הסכמה בין שופטים, שני אנשים יכולים לקבל ציונים שונים מסיבות שונות



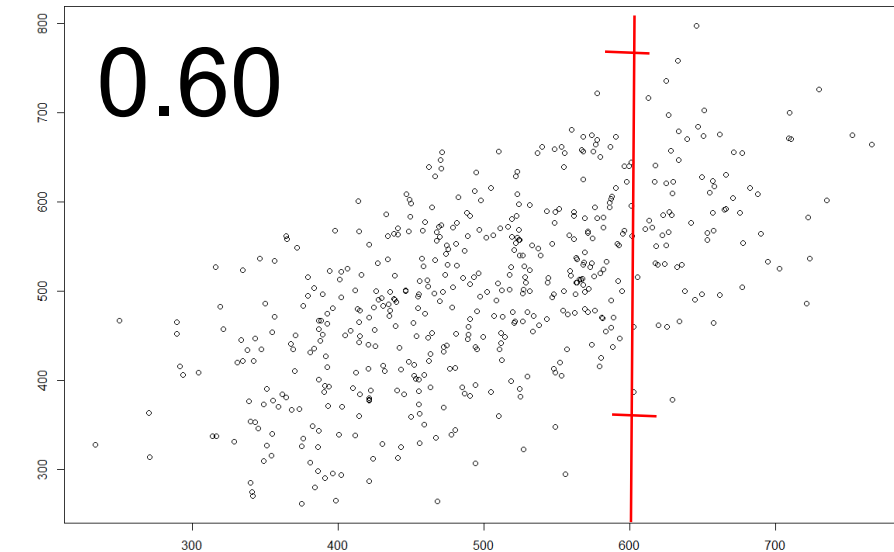
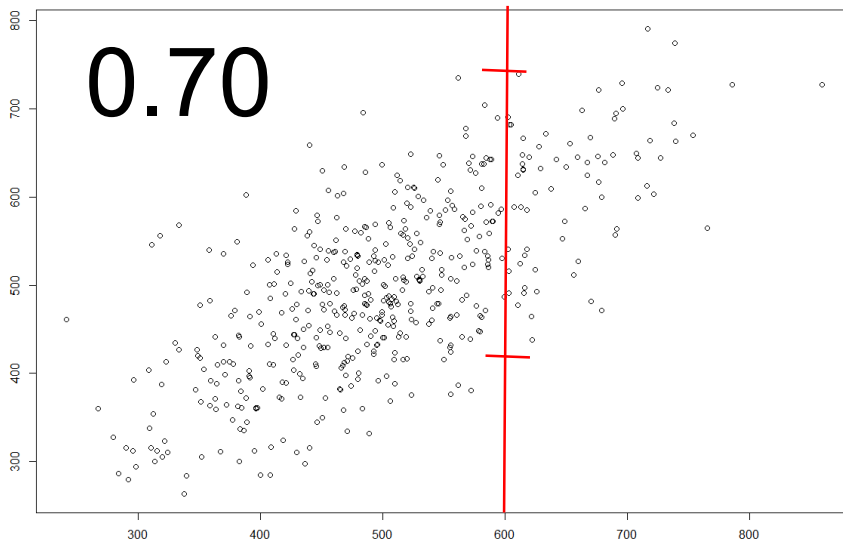
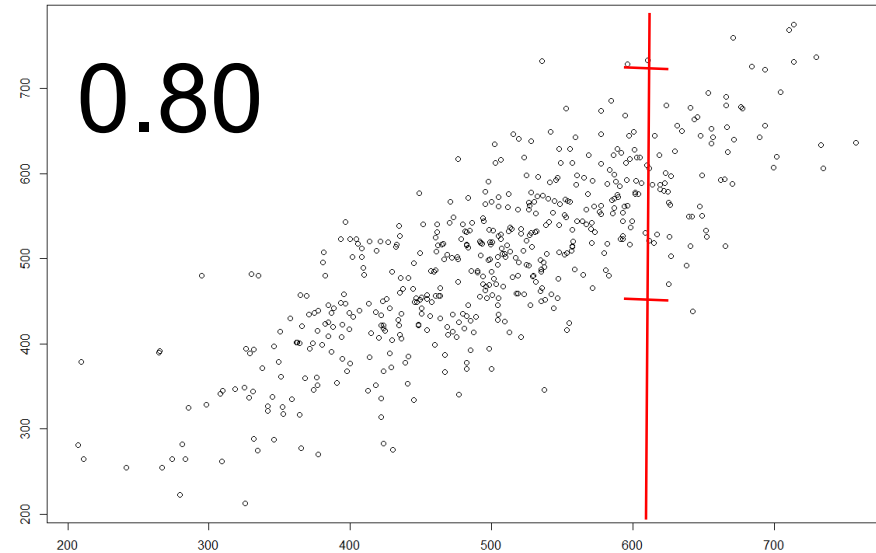
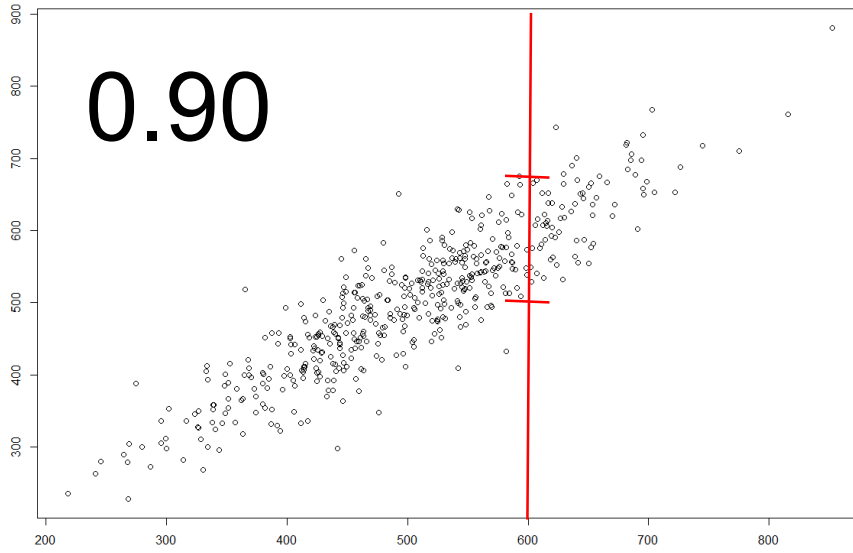
טעות אמדן-דגימה (sampling adequacy)

- מבחן מבוסס על דגימה של התנהגויות / מדידות
 - במבחן יידע, לא שואלים על כל החומר
 - במבחן חרדה לא שואלים על כל סוג מחשבה או התנהגות הקשורה בחרדה
- התוצאה של המדגם של ההתנהגויות של האדם לא משקפת בדיוק את התוצאה באוכלוסית כל ההתנהגויות הרלוונטיות של האדם ("**הציון האמיתי**")
 - אם היינו שואלים על כל החומר
 - אם היינו שואלים על כל היבט התנהגותי הקשור לחרדה

שיטות להערכת איכות דגימה-אמדן (sampling adequacy)

- מהימנות של נוסחים מקבילים
- מתאם בין ציונים שהתקבלו בנוסחים שונים של המבחן
- מהימנות חצויה (פנימית)
- מתאם בין ציונים שחושבו על שתי מחציות של הפריטים
- נאמר זוגיים מול אי זוגיים

מהימנות פנימית קובעת את מידת ההחטאה באמדן



מידת ההחטאה – טעות התקן של האמדן קשורה לאיכות הדגימה – מהימנות פנימית:

- Standard error of measurement = SEM

- $SEM = SD\sqrt{1 - Reliability}$

- הדגמה במונחים של ציון פסיכומטרי (סטיית תקן=100)

- מהימנות פנימית 0.90, SEM=32, 568-632

- מהימנות פנימית 0.80, SEM=45, 555-645

- מהימנות פנימית 0.70, SEM=55, 545-655

- מהימנות פנימית 0.60, SEM=63, 537-663

- חוסר מהימנות פנימית קריטי במיוחד כשבודקים אדם בודד
 - באבחון ובמיון

- פחות קריטי במחקר על קבוצת אנשים
 - ישנן דרכים סטטיסטיות לטפל בחוסר מהימנות

מה משפיע על מהימנות פנימית

- מידת ההטרוגניות של התכונה וגודל המדגם

- אנאלוגי בדיוק לדגימה בכלל

הטרוגניות



σ

גודל מדגם



\sqrt{N}

- טעות התקן של הממוצע במדגם

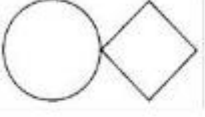







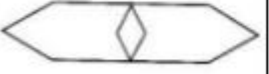
גם אם הדגימה טובה (מהימנות פנימית גבוהה) ייתכן שהמבחן בודק מצב ולא תכונה

- זו אינה בעיה במבחן אלא ממצא של מחקר
- יש לתוצאה כזו השלכות
- לא ניתן לנבא (טוב) התנהגויות עתידיות
- אם הציון לא מנבא את עצמו הוא לא ינבא דברים אחרים
- מבחן כזה מתאים במיוחד למדידת השפעות מצביות (תוצאות טיפול, תוצאות אימון, מצב רוח)

תוקף – מה המבחן בודק? מה משקף ציון גבוה/נמוך במבחן?

- ישנם סוגים שונים של תוקף:
 - תוקף נראה Face Validity
 - תוקף תוכן Content Validity
 - תוקף ניבוי Predictive Validity
 - תוקף בו-זמני Concurrent Validity
 - תוקף מבנה Construct Validity

תוקף נראה – מה נראה כי המבחן בודק?

5. 	6. 	7. 	8. 
9. 	10. 	11. 	12. 
13. 			

• מבחן העתקת צורות

- למעשה המבחן אמור לבדוק אישיות
- האם התוקף הנראה גבוה?
- האם טוב שכך?

• מה בודקת התשובה לשאלה "אני אדם יציב"?

- האם התוקף הנראה גבוה?
- האם טוב שכך?

• ישנם מצבים שבהם התוקף הנראה נמוך לנבדק וגבוה לבודק

- מבחן פריימינג

- לנבדק – בודק יכולת קריאה
- לבודק – בודק זיכרון

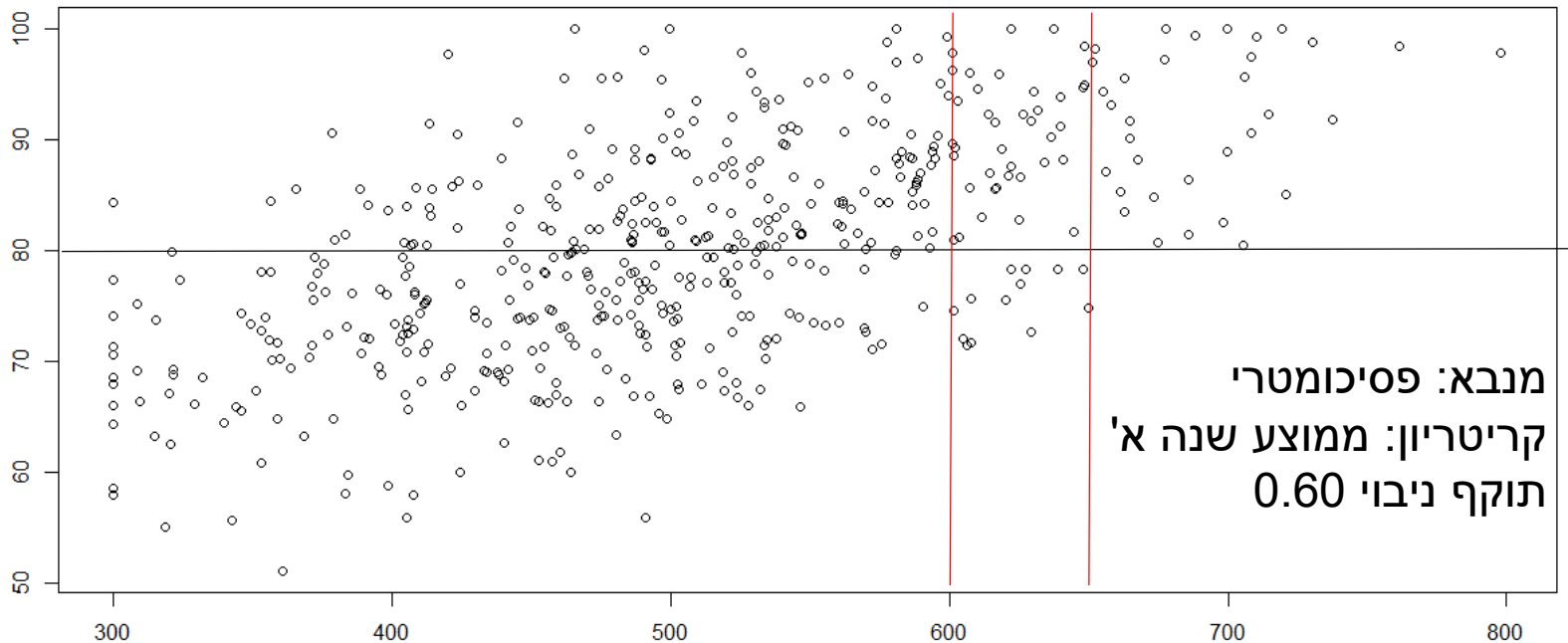
תוקף תוכן

- האם כל התכנים הרלוונטים מיוצגים במבחן?
- האם הייצוג שלהם פרופורציונלי?
- במבחן יידע
 - כיסוי בפרופורציה מתאימה לכל הפרקים
- במבחן שבודק תכונה
 - כיסוי לכל האספקטים הרלוונטיים של התכונה
- דוגמה: מבחן MSCEIT של אינטליגנציה רגשית בודק את ארבעת התכנים הרלוונטיים על פי התיאוריה:
 - זיהוי רגשות
 - שימוש ברגשות
 - הבנת רגשות
 - ידע לגבי ויסות רגשות

תוקף תלוי קריטריון

- המתאם הסטטיסטי בין הציון במבחן לבין הציון בקריטריון
 - אם הקריטריון הוא הצלחה בתפקיד: **תוקף ניבוי**
 - אם הקריטריון הוא מדד אחר בעל משמעות תיאורטית: **תוקף מתכנס, תוקף בו זמני**
 - מתאם בין ציון באינטליגנציה לסטטוס סוציו-אקונומי
 - מתאם בין אינטליגנציה רגשית לבין פופולריות בבית ספר

מדוע אפשר להשתמש במבחנים עם תוקף ניבוי לא גבוה במיוחד לצורך מיון?



• מיון מוריד False Alarms - מתקבלים שלא יקבלו ציון 80 ומעלה

• 50.6% השיגו ממוצע מעל ל 80

• אם מקבלים רק מפסיכומטרי מעל 600, 84.5%

• אם החתך עולה ל 650, אחוז זה עולה ל 100%.

• המחיר: Misses - נדחים שהיו מקבלים ציון 80 ומעלה

• ללא חתך אין החטאות

• שיעור ההחטאות הוא 43.8% בחתך של 600, והוא עולה ל 47.1% בחתך 650.

תוקף מבנה – Construct Validity

- כל המידע שיש לנו אודות המבחן שמעיד על מה המבחן בודק
- **מהימנות חוזרת:** האם המבחן בודק תכונה יציבה או מצב?
- **תוקף מתכנס (מיצוי):** מתאם גבוה יחסית בין ציון במבחן לציון במבחנים שבודקים תכנים קרובים או רלוונטיים:
 - מתאם גבוה בין אינטליגנציה רגשית לפופולריות בבית ספר
- **תוקף מבחין (בלבדיות):** מתאם נמוך יחסית עם ציונים במבחנים שבודקים דברים אחרים
 - מתאם נמוך בין אינטליגנציה רגשית לאישיות

שני אנשים יכולים לקבל ציונים שונים מסיבות שונות (דוגמה לגבי מבחן אינטליגנציה)

• בעיית אמדן

- לשניהם אותה אינטליגנציה, אבל לאחד מהם היא מתבטאת ביכולות מרחביות ואצל האחר היא מתבטאת ביכולות כמותיות. המבחן בודק רק אינטליגנציה כמותית.

• מצב

- אחד מהם היה במקרה ביום גרוע

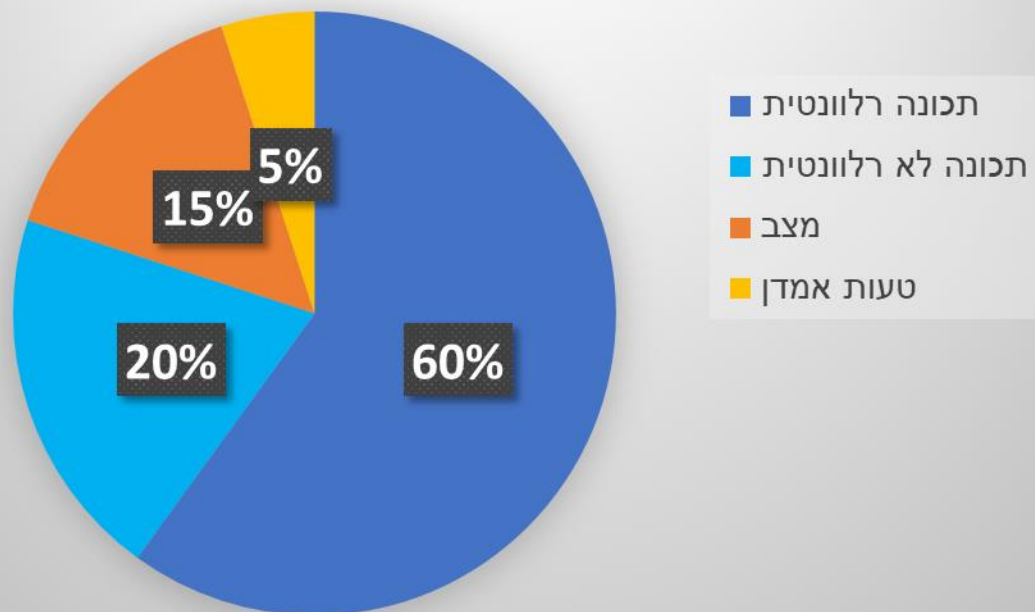
• תכונה לא רלוונטית

- לשניהם אינטליגנציה זהה, אבל לאחד מהם יש יותר מוטיבציה להצליח במבחן

• תכונה רלוונטית

- יש ביניהם הבדלי אינטליגנציה

הבדלים בינאישיים בתוצאות
המבחן



נורמות: כיצד יודעים אם ציון הוא גבוה או נמוך?

- האם ציון 20 ב MMSE הוא ציון גבוה?

- טווח הערכים האפשרי 0-30

- החתך הקליני הוא 24

- **הנורמה היא החתך הקליני**

- נורמות המבוססות על ציוני תקן

- בכמה סוטה מהממוצע?

- הסטייה מהממוצע מבוטאת ביחידות הקשורות לסטיית התקן

- פסיכומטרי 650 נמצא 1.5 סטיות תקן מהממוצע

- **הנורמה היא הממוצע וסטיית התקן**

- מקובל לקבוע ממוצע וסטיית תקן

- IQ – ממוצע 100, ס"ת 15

- פסיכומטרי – ממוצע 500, ס"ת 100

- נורמות הפתחותיות

- לאיזה גיל מתאים ציון כזה?

- ילדה בת 8 המבצעת כמו בני 12

- **הנורמה היא הממוצע (או החציון) לפי גיל**

- $IQ = (\text{Mental age} / \text{Chronological age}) * 100$

- לילדה בדוגמה $IQ = (12/8) * 100 = 150$