

מבחנים וחברה

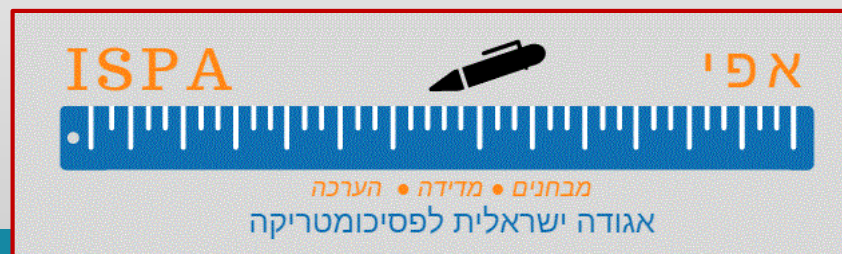
פאנל מומחים - שאלות ותשובות, דיון עם הקהל

מארגן ומנחה: אבי אללוף

משתתפים: עינב אבני, ענת בן סימון, גרשון בן שחר, אייל גמליאל, רענן הס

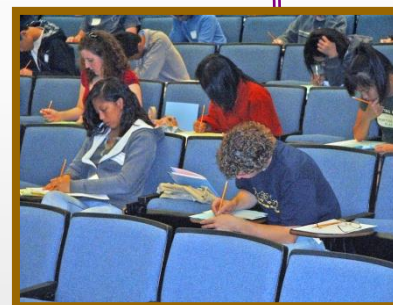
הזום

כנס אפי, מעלה החמישה, 2022



מבחנים משפיעים עלינו...

לאורך החיים



מבחנים – חינוכיים, פסיכולוגיים ומקצועיים - משמשים להחלטות חשובות בתחומים רבים, כמו

- **קבלה לעבודה וללימודים גבוהים**
- **אבחון אישיותי לפני קבלת טיפול פסיכולוגי**
- **בדיקת אמינות ע"י הפוליגרף**
- **קביעת מדיניות חינוך על סמך הישגים במבחנים ארציים סטנדרטיים**
- **קביעת קריטריונים להעדפה מתקנת.**

מבחנים – חינוכיים, פסיכולוגיים ומקצועיים - משמשים להחלטות חשובות בתחומים רבים, כמו

- **קבלה לעבודה וללימודים גבוהים**
 - **אבחון אישיותי לפני קבלת טיפול פסיכולוגי**
 - **בדיקת אמינות ע"י הפוליגרף**
 - **קביעת מדיניות חינוך על סמך הישגים במבחנים ארציים סטנדרטיים**
 - **קביעת קריטריונים להעדפה מתקנת.**
-
- **האם ההחלטות שנלקחות על סמך המבחנים הן נכונות?**
 - **האם השימוש הרב המתבצע כיום במבחנים הוא מוצדק?**
 - **האם מבחני ההסמכה משיגים את מטרתם, ולחילופין – האם יש תחומים שבהם מבחני הסמכה חסרים?**
 - **האם המבחנים מנציחים פערים בחברה? האם הם הוגנים?**
 - **האם נפוצות ההכנה למבחנים פוגעת בתהליכי מיון?**
 - **האם חשוב לחזק את האוריינות בציבור בנושאי מדידה והערכה?**
 - **כיצד ייראו המבחנים בעתיד?**

מבחנים – חינוכיים, פסיכולוגיים ומקצועיים - משמשים להחלטות חשובות בתחומים רבים, כמו

- **קבלה לעבודה וללימודים גבוהים**
- **אבחון אישיותי לפני קבלת טיפול פסיכולוגי**
- **בדיקת אמינות ע"י הפוליגרף**
- **קביעת מדיניות חינוך על סמך הישגים במבחנים ארציים סטנדרטיים**
- **קביעת קריטריונים להעדפה מתקנת.**

- **האם ההחלטות שנלקחות על סמך המבחנים הן נכונות?**
- **האם השימוש הרב המתבצע כיום במבחנים הוא מוצדק?**
- **האם מבחני ההסמכה משיגים את מטרתם, ולחילופין – האם יש תחומים שבהם מבחני הסמכה חסרים?**
- **האם המבחנים מנציחים פערים בחברה? האם הם הוגנים?**
- **האם נפוצות ההכנה למבחנים פוגעת בתהליכי מיון?**
- **האם חשוב לחזק את האוריינות בציבור בנושאי מדידה והערכה?**
- **כיצד ייראו המבחנים בעתיד?**

אלה הן חלק מהשאלות הרבות הממחישות עד כמה המבחנים משפיעים על החברה כמו גם על הפרט, כאשר במקביל – החברה משפיעה על המבחנים ומחליטה איזה משקל יינתן להם בהחלטות חשובות.

מבחנים וחברה

- חשוב שאנשי המקצוע יבינו את ההשפעה שיש למבחנים על החברה מעבר לידע שלהם במתודולוגיה
- חשוב גם שסטודנטים הלומדים מדידה והערכה יכירו את הסוגיות העיקריות
- חשוב שקובעי מדיניות, תקשורת ואף כלל הציבור יהיו מודעים לכך

הקורס שאני מלמד מהווה גם חלק מלימודי התעודה בפסיכומטריקה



פאנל של חמישה משתתפים בעלי מומחיות במדידה והערכה, הבאים מרקע שונה, מהאקדמיה, מארגונים ציבוריים וחברות פרטיות, ישיבו על שאלות אחידות, מתוכננות מראש.

השאלות סוכמו מראש עם המשתתפים מתוך מטרה לאפשר הצגת גישות שונות ומגוון דעות והשקפות מבוססות נתונים.

יוקצה זמן מוקצב לתשובה של כל אחד מחברי הפאנל.

3 דקות לתשובה

שאלה	משיבים
מהי השפעת ההכנה על המבחנים?	עינב
	גרשון
	אייל
מה דעתך על מדיה חברתית/טביעת רגל דיגיטלית בדרך להעריך אנשים?	ענת
	רענן
	עינב
מהי התרומה הכלכלית של המבחנים?	גרשון
	רענן
האם הציבור יכול להעריך את איכות המבחנים?	ענת
מה הקשר שבין תוקף והוגנות של מבחנים?	אייל
האם יש מבחנים המזיקים לחברה?	גרשון
כיצד ייראו המבחנים בעתיד?	רענן
	עינב
מה האתגר הגדול ביותר שנתקלת בו בעבודתך בעבר או בעבודתך כיום?	אייל
	ענת



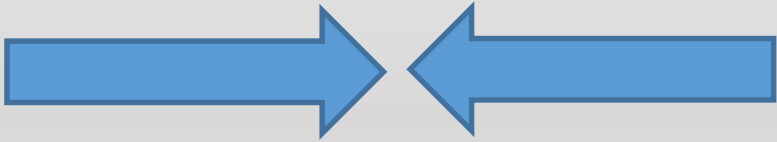
חישוב זמנים.....70 דקות

- הקדמה..... **5** דקות
- 8 שאלות, 0.5 דקה לשאלה..... **4** דקות
- 15 תשובות 3 דקות לתשובה..... **45** דקות
- סה"כ..... **54** דקות

- זמן שנותר לדין אחרי הפאנל..... **16** דקות



חברה



מבחנים

שאלה 1: מהי השפעת ההכנה על המבחנים?

עינב אבני, גרשון בן שחר, אייל גמליאל

חסרונות

- פגיעה בסטנדרטיזציה
- פגיעה ביכולת הניבוי
- פגיעה בהערכת יכולת התמודדות בלחץ

יתרונות

- "סינון החוצה"
- שיפור תוקף המיון
- תחושת הוגנות ושקיפות
- שיוויון הזדמנויות

סוגים של הכנות:

מידע עובדתי
על המיון

הצגת כלי
ותוכן המיון

מתן כלים/שיטות
לביצוע המיון

מהי השפעת ההכנה על המבחנים?

- Misconception about the nature of aptitude tests.
- The study conducted by Avi Allalouf (Allalouf & Ben-Shakhar, 1998).
- While coaching increased PET scores by $.25Sd$, it did not affect test validity and did not create bias against the un-coached group.

מהי השפעת ההכנה על המבחנים?

מבחני הישגים לעומת "מבחנים קוגניטיביים".

אתגרי מבחני האיתור למחוננים.

התנסות אישית.

אפקטיביות ההכנה.

Allalouf & Ben-Shakhar (1998): 0.25 SD
(preparation/coaching on PET)

Hausknecht et al., 2007: K=107; N=135,000;
0.26 SD (practice effects -- cognitive ability tests)

רבע סטיית תקן זהו הפער בין אחוזון 50 ל 60 בהתפלגות נורמלית.

שאלה 2: מה דעתך על מדיה חברתית/טביעת רגל דיגיטלית כדרך להעריך אנשים?

ענת בן סימון, רענן הס, עינב אבני

שאלה 2: מה דעתך על מדיה חברתית/טביעת רגל דיגיטלית כדרך להעריך אנשים? רענן הס A

מה דעתך על מדיה חברתית/טביעת רגל דיגיטלית כדרך להעריך אנשים

- השוק מעריך את השימוש בטכנולוגיה יותר מאשר את הפסיכולוגיה
- זו כמובן התחרות הקשה ביותר שלנו, אם כי לא מעט אנשים מדברים על כך שנעשה Abuse של המילה
- נכון להיום האוטומציה מתחזקת על חשבון הפסיכולוגיה. בהייטק חוששים לגעת במועמדים
- AI/BI מתאימים למיון גנרי במאסות
- WEB 3.0: Deep Learning, ביטקוין, זהות דיגיטלית- אווטאר

שאלה 2: מה דעתך על מדיה חברתית/טביעת רגל דיגיטלית כדרך להעריך אנשים? רענן הס B

מה דעתך על מדיה חברתית/טביעת רגל דיגיטלית כדרך להעריך אנשים

- לכלי AI יש אולי יכולת כאשר מדובר בתפקיד ספציפי לאורך זמן, ויש צורך בדטה!!
- לדעתי, לכלי AI אין את היכולת לנבא ברמת דיוק של תפקיד בארגון ספציפי, תחת סגנון ניהול ספציפי ותחת תרבות ארגונית ספציפית – לנו יש (אם נשתנה)
- כדי לנבא עם AI יש צורך בדטה איכותי, איסוף של דטה איכותי לוקח זמן ודורש מאסות. לתפקידים בכירים אין מסות
- התבקשתי למיין לתפקיד סמנכ"ל שיווק לחברה מסויימת
- מומחה הציע לי NLP, אבל אנחנו יודעים שהיום אנשים משתמשים במילים כשהם שולחים קורות חיים
- מומחה הציע לי לעשות שימוש ב-AI, אבל הבין שאין לי דטה. אז הציע שאפנה לארגונים בינלאומיים לקבל מהם דטה. אבל אז אמר מדוע שיתנו לך?
- הפסיכולוגיה מאפשרת לנו למיין מהר יותר ובדיוק גבוה יותר מכל כלי אחר, אם נעבור מ-Skills based approach ל-Performance based approach. דהיינו למיין תפור לתפקיד
- שימוש בבטריה מעט משתנה מתוך מאגר המבחנים הפסיכולוגיים שלנו לא עושה הבדל – אנחנו כמו AI

שאלה 2: מה דעתך על מדיה חברתית/טביעת רגל דיגיטלית כדרך להעריך אנשים? איננו אבני

מדיה חברתית ומשמעותה על הערכה



שאלה 3: מהי התרומה הכלכלית של המבחנים?

גרשון בן שחר, רענן הס

מהי התרומה הכלכלית של המבחנים?

How valid a test should be to justify its use? Validity vs. utility

Several models have been developed to translate Test's validity into utility estimates. For example, Schmidt et al. (1979) applied such a model and demonstrated that even low validity (e.g., 0.2) may have great impact on work force productivity and the economy.

Admission tests in a quota free context (Ben-Shakhar & Beller. 1983; Ben-Shakhar et al., 1996). Using The combined predictor is justified from a cost/benefit perspective. The marginal utility of PET is positive if the goal is excellence.

מהי התרומה הכלכלית של המבחנים?

- צריך לחשוב במונחים כלכליים, אנחנו צריכים לראות את עצמנו כאנשי "רכש הון אנושי"!
- איש רכש מקצוען לא מסתפק בלקבל את הדרישות הוא מבקש לדעת את המטרות/את הציפיות מהרכש.
- יש להביא לארגון את המועמד המתאים ביותר שייתן לארגון ערך מהיום הראשון.
- לכן צריך להבין את הציפיות מהעובד החדש מבחינת **הערך** שהעובד אמור לתת לארגון, מה הציפיות ממנו?

- אין למבחנים הפסיכולוגים, נכון להיום, תרומה כלכלית כי תרומה שכזו לא מוגדרת!
- אני חושב שאנשי משאבי אנוש לא חושבים במונחים כלכליים ולכן יש מידה די גדולה של חוסר הערכה בארגון כלפיהם
- הציפיה מאתנו היא לנבא את הסיכוי לקבל ערך (TTV) מהעובד החדש.
- כאזר מנבאים יכולות אי אפשר לתת ערך כלכלי, אפשר להוריד סיכון!

- המחיר הכלכלי של תוקף ניבוי נמוך
- תחלופת עובר עולה כ- 6 משכורות
- מחסור בכ"א לא מאפשר לארגון לעמוד ביעדים שלו

שאלה 4: האם הציבור יכול להעריך את איכות המבחנים?

ענת בן סימון

האם הציבור יכול להעריך את איכות המבחנים?

- שוויון הזדמנויות
- פערים בציוני אוכלוסיות
- שיטות להערכה הטיה (תיאורטיות/אמפיריות)
- התאמות בבחינות

הוגנות

- תוקף נראה
- תוקף תוכן
- תוקף חיזוי
- תוקף אקולוגי
- ...

תוקף

- מהימנות עקיבות פנימית
- מהימנות מבחן חוזר
- מהימנות בין שופטים
- מדדי אמינות
- ...

מהימנות

רמות ביצוע

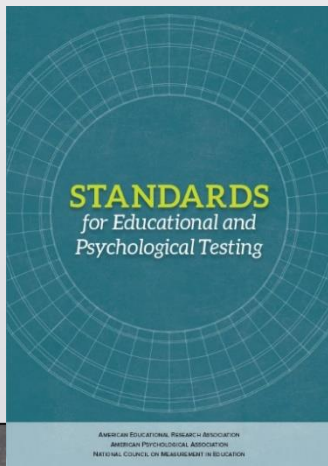
IRT

תקנים

נורמות

כיול
(אורך רחב)

סילום



האם הציבור יכול להעריך את איכות המבחנים?



- ❖ **לא!** לציבור או למקבלי החלטות "הדיוטות" (לא מומחים) אין שום יכולת (או רצון) לבחון את איכות הכלים שהם עושים בהם שימוש
- ❖ לרוב הבחירה בכלי זה או אחר נגזרת משיקולים שוליים (זמינות ועלות)
- ❖ הדרך היחידה להעריך איכות מבחן היא לתת את אותו ואת כל המידע שנצבר עליו לשיפוט של מומחים בתחום המבחנים

שאלה 5: מה הקשר שבין תוקף והוגנות של מבחנים?

אייל גמליאל

מה הקשר בין תוקף והוגנות של מבחנים? האם מחקרי התוקף משתמשים בקריטריון הרלוונטי?

מבחני מיון, מבחני הסמכה, מבחני קבלה, ומבחנים אחרים פועלים במסגרת חברתית.

לרוב, מחקרי התיקוף משתמשים בקריטריון הצלחה אינדיווידואלי (ממוצע לימודים בתואר; הערכות ממונים).

האם לא צריך/כדאי לשקול את ההקשר החברתי של השימוש במבחנים ולהגדיר את קריטריון הצלחה באופן שיתייחס להקשר החברתי והתרבותי שלשמו משתמשים במבחן.

יותר ויותר מקומות עבודה פועלים בהתאם לעיקרון:
גיוון והכלה -- Diversity & Inclusion.

שאלה 6: האם יש מבחנים המזיקים לחברה?

גרשון בן שחר

האם יש מבחנים המזיקים לחברה?

- Throughout my career I have tried to fight (usually unsuccessfully) applications based on pseudoscience.
- Among those, I can mention 2 tests that are being used extensively for various purposes.
- Lie-detection, based on the so called Control Question Test (CQT), used in forensic investigations, but also for selecting and screening candidates.
- Hand-writing analysis (graphology), used for personnel selection.

שאלה 7: כיצד ייראו המבחנים בעתיד?

רענן הס, עינב אבני

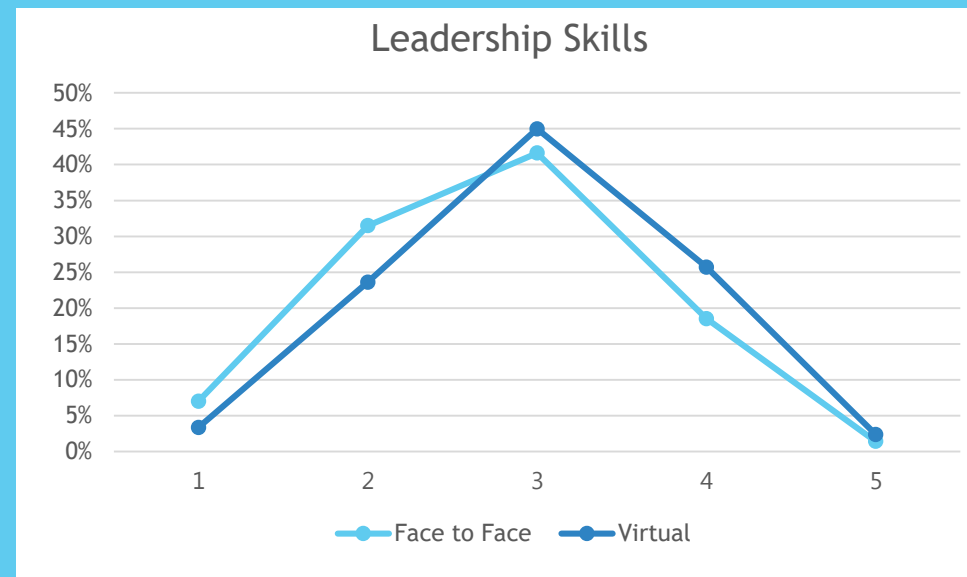
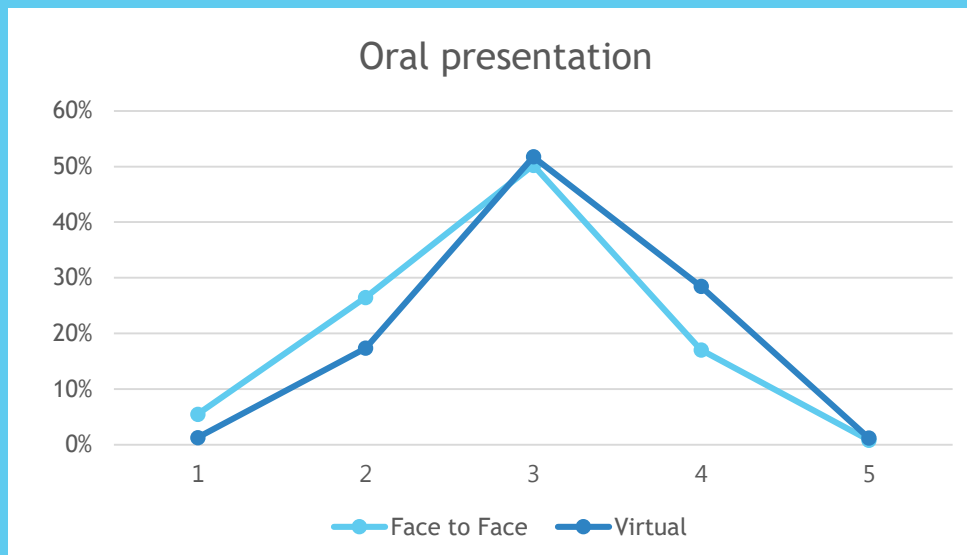
כיצד ייראו המבחנים בעתיד?

בעתיד:

- אלגוריתם שייקח בחשבון פסיכולוגיה ו- AI
- יהיה לא רק שימוש בגישה - Skills based approach, אלא שימוש בדגש על Performance based approach. דהיינו ננסה לנבא תוצאות ולא רק יכולות
- אנחנו עוסקים היום בניבוי תפיסת תפקיד
- המבחנים הפסיכולוגיים המסורתיים יהוו כלים שוללים, מפחיתי סיכון
- ואולי יוחלפו באפליקציות מסוג: VR, Gaming וכדומה

שאלה 7: כיצד ייראו המבחנים בעתיד? עינב אבני

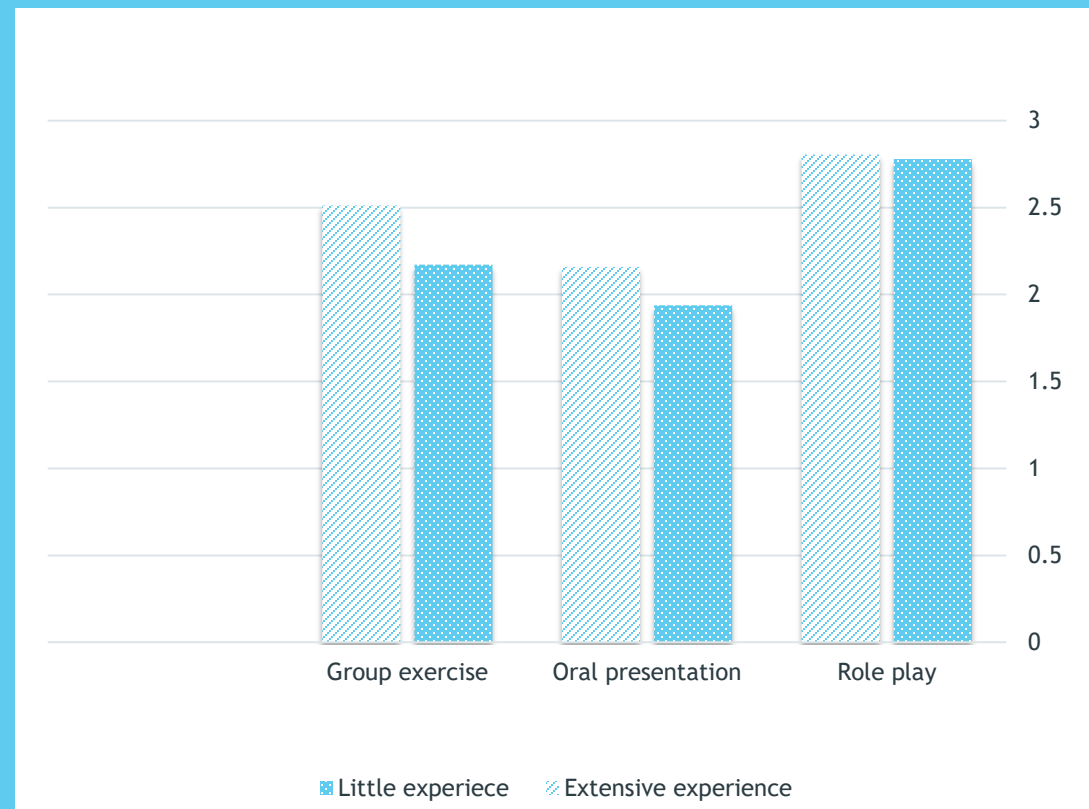
התפלגות הערכות פרונטאלי מול וירטואלי



כלי מיון עתידיים מרכז הערכה וירטואלי מול מרכז הערכה פרונטאלי



תחושת הביטחון של המעריכים



שאלה 8: מה האתגר הגדול ביותר שנתקלת בו בעבודתך בעבר או בעבודתך כיום?

אייל גמליאל, ענת בן סימון

שאלה 8: מה האתגר הגדול ביותר שנתקלת בו בעבודתך בעבר או בעבודתך כיום? אייל גמליאל A

מה האתגר הגדול ביותר שנתקלת בו בעבודתך בעבר או בעבודתך כיום?

מדוע מבחן חסר הטיה בניבוי הוא בהכרח מבחן שמפלה בברירה
לרעה קבוצות מוחלשות?

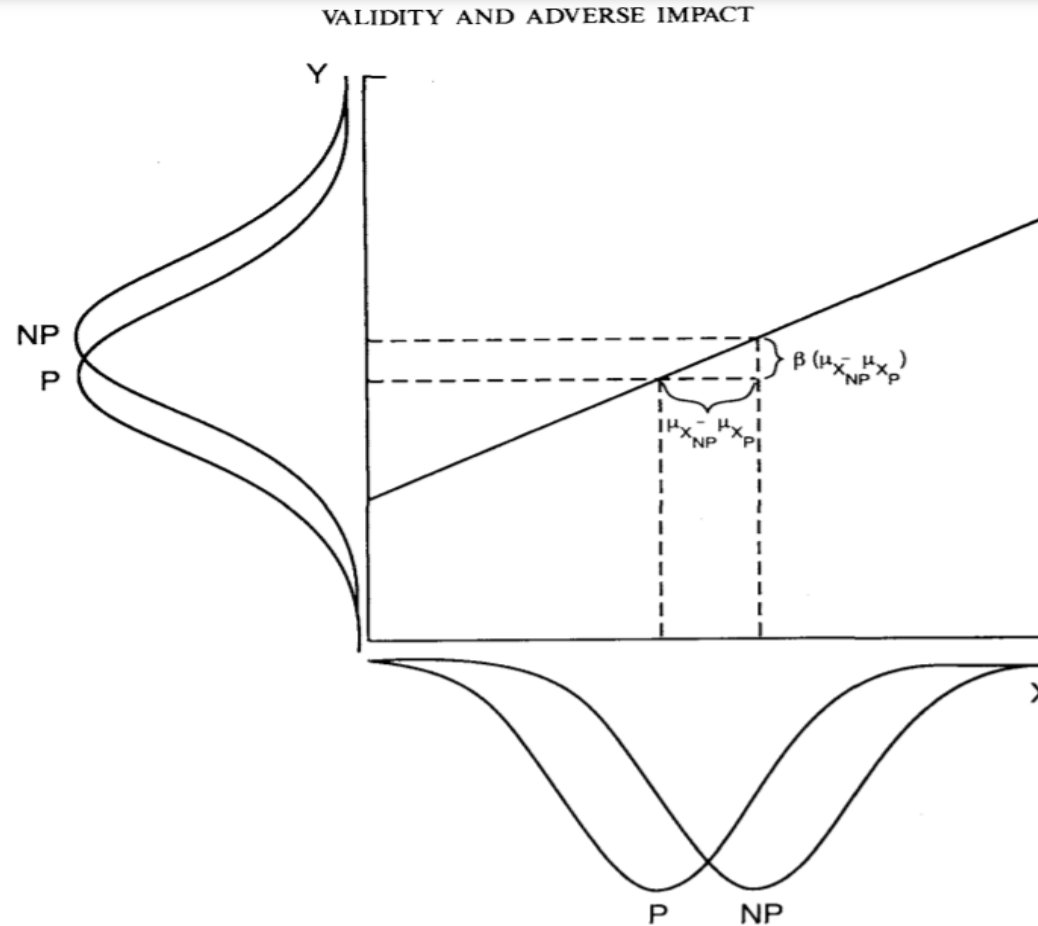


Figure 1. Relation between criterion (Y) and a fair predictor (X) in two populations (P and NP).

מה האתגר הגדול ביותר שנתקלת בו בעבודתך בעבר או בעבודתך כיום?

Robyn Dawes (2002):

The overlap on the criterion between minority and majority groups will be greater than the overlap on the predictors.

If we use the same statistical prediction rule for minority and majority group members, we're differentially rejecting a higher proportion of minority group members who would have done well than of majority group members who would have done well.

שאלה 8: מה האתגר הגדול ביותר שנתקלת בו בעבודתך בעבר או בעבודתך כיום? ענת בן סימון



מה האתגר הגדול ביותר שנתקלת בו בעבודתך בעבר או בעבודתך כיום?



לבחון 40,000 מועמדים במועד נתון לפי הנחיות משרד הבריאות שפורסמו ערב קודם

ובימים כתיקנם –

- ❖ ליישב את המתח בין איכות כלי ההערכה, עלות הפיתוח שלהם ומחירם למשתמש
- ❖ לתקשר את המשמעות והתפקיד של הבחינה הפסיכומטרית ונושא ההתאמות בבחינה לפוליטיקאים, לתקשורת ולציבור הרחב