

יום עיון בכתיבת מבחנים

יום עיון בכתובת מבחנים

סדר יום:

10:00-10:40 היכרות עם תורת המבחנים

10:40-12:00 קווים מנחים לכתובת שאלות בררה

12:00-12:15 הפסקה

12:15-13:15 חיבור שאלות פתוחות והערכתן

13:15-14:00 איך בונים מבחן

היכרות עם תורת המבחנים

מהו מבחן?

✓ הליך סטנדרטי של מדידת ידע המבוסס על דגימה של תחום הידע ומשמש להסקת מסקנות ו/או קבלת החלטות.

מטרות שונות למבחנים

- ✓ אבחון קליני / דיאגנוסטיקה
- ✓ הערכת ידע
- ✓ עידוד למידה
- ✓ מיון: פסיכומטרי
- ✓ סינון: יחידות בצבא
- ✓ רישוי / הסמכה: רפואה, עריכת דין
- ✓ ייעוץ והכוון

סוגים של הערכה

✓ הערכה מסכמת (Summative Evaluation)

נעשית בסוף תהליך הלמידה, ומעריכה את הישגיו של התלמיד.

✓ הערכה מעצבת (Formative Evaluation)

הערכה שנעשית כחלק מתהליך הלמידה, ומטרתה לתת משוב לסטודנטים, לסייע להם ולמרצה להשתפר.

סוגים של מבחנים

✓ מבחני הישגים – מבחנים המודדים שליטה בתחום ידע מוגדר. למשל, מבחן שניתן בסוף קורס ובוחן תכנים שנלמדו בו.

✓ מבחני יכולת – מבחנים שאינם מודדים ידע מוגדר, אלא יכולת מולדת או יכולת שהתפתחה לאורך זמן בתרגול ובהתנסות. למשל, מבחני אינטליגנציה.

← מבחני יכולת משמשים לעיתים קרובות למיון, כלומר לחיזוי הישגיו של הנבחן במצב חדש.

סוגים של מבחנים

✓ מבחני נורמה - מבחנים שמטרתם להעריך את ההישגים של הנבחן לעומת קבוצת נבחנים שקרויה "קבוצת הנורמה". ציון במבחן נורמה מציין את מיקומו היחסי של הנבחן בקבוצה זו.

✓ מבחני קריטריון - מבחנים שמטרתם להעריך את ההישגים של הנבחן לעומת עולם תוכן מסוים.

טעויות בהערכה

✓ בכל מדידה טבוע מרכיב טעות.

✓ סוגי הטעויות המוכרים מעולם המחקר שימושיים גם בתורת המבחנים:

טעות מסוג א - אדם בלתי-כשיר שמוכרז כשיר.

טעות מסוג ב - אדם כשיר שמוכרז בלתי-כשיר.

נקודת החתך (במבחני קריטריון)

נקודת החתך היא הציון שמעליו האדם ייחשב שולט בחומר שליטה מספקת.

לא צריך לחזור על הקורס	ציון "עובר"
נקודת החתך	
צריך לחזור על הקורס	ציון "נכשל"

✓ ככל שנקודת החתך נמוכה יותר, כך גדלה ההסתברות לטעות מסוג ראשון.

✓ ככל שנקודת החתך גבוהה יותר, כך גדלה ההסתברות לטעות מסוג שני.

ביקורת על מבחנים

- ✓ שיפוט משתנים איכותיים בקריטריונים כמותיים
- ✓ חוסר רגישות ליצירתיות ולמקוריות
- ✓ שימוש במבחן למטרות שאינן מתאים להן
- ✓ הטיה חברתית ושמירה על ההגמוניה של אליטות חברתיות
- ✓ עידוד "לא ערכי" ללמידה ועידוד תחרותיות
- ✓ "חיפוש מתחת לפנס"

פריטים סגורים ופתוחים

מה ההבדל בין פריטים סגורים לפריטים פתוחים?

פריטים סגורים -

המיומנות: זיהוי (recognition)

פריטים פתוחים -

המיומנות: שליפה (recall)

השוואה בין פריטים סגורים לפריטים פתוחים

פריטים פתוחים	פריטים סגורים	
יצור התשובה הנכונה	איתור התשובה הנכונה	הדרישות מהנבחן

השוואה בין פריטים סגורים לפריטים פתוחים

פריטים פתוחים	פריטים סגורים	
ייצור התשובה הנכונה	איתור התשובה הנכונה	הדרישות מהנבחן
גמיש - השאלות לרוב כלליות	מצומצם - השאלות ממוקדות	היקף החומר בפריט מסוים

השוואה בין פריטים סגורים לפריטים פתוחים

פריטים פתוחים	פריטים סגורים	
ייצור התשובה הנכונה	איתור התשובה הנכונה	הדרישות מהנבחן
גמיש - השאלות לרוב כלליות	מצומצם - השאלות ממוקדות	היקף החומר בפריט מסוים
ארוך	קצר למדי	משך הפתרון של פריט מסוים

השוואה בין פריטים סגורים לפריטים פתוחים

פריטים פתוחים	פריטים סגורים	
ייצור התשובה הנכונה	איתור התשובה הנכונה	הדרישות מהנבחן
גמיש - השאלות לרוב כלליות	מצומצם - השאלות ממוקדות	היקף החומר בפריט מסוים
ארוך	קצר למדי	משך הפתרון של פריט מסוים
קטן	גדול	מספר השאלות במבחן כולו

השוואה בין פריטים סגורים לפריטים פתוחים

פריטים פתוחים	פריטים סגורים	
ייצור התשובה הנכונה	איתור התשובה הנכונה	הדרישות מהנבחן
גמיש - השאלות לרוב כלליות	מצומצם - השאלות ממוקדות	היקף החומר בפריט מסוים
ארוך	קצר למדי	משך הפתרון של פריט מסוים
קטן	גדול	מספר השאלות במבחן כולו
יכולה להיות גבוהה	נמוכה	איכות המשוב

השוואה בין פריטים סגורים לפריטים פתוחים

פריטים פתוחים	פריטים סגורים	
ידע, הבנה ויישום, וכן חשיבה מקורית, הבעה, יכולת כתיבה וכדומה	ידע, הבנה, יישום וכדומה	מהן המיומנויות שאפשר לבדוק?

השוואה בין פריטים סגורים לפריטים פתוחים

פריטים פתוחים	פריטים סגורים	
ידע, הבנה ויישום, וכן חשיבה מקורית, הבעה, יכולת כתיבה וכדומה	ידע, הבנה, יישום וכדומה	מהן המיומנויות שאפשר לבדוק?
בבניית המחווון ובבדיקת הפריטים	בחיבור הפריטים	עיקר המאמץ

השוואה בין פריטים סגורים לפריטים פתוחים

פריטים פתוחים	פריטים סגורים	
ידע, הבנה ויישום, וכן חשיבה מקורית, הבעה, יכולת כתיבה וכדומה	ידע, הבנה, יישום וכדומה	מהן המיומנויות שאפשר לבדוק?
בבניית המחווון ובבדיקת הפריטים	בחיבור הפריטים	עיקר המאמץ
דורשת מחשבה ומומחיות והיא סובייקטיבית	קלה, מהירה, זולה ואובייקטיבית	בדיקת השאלה

השוואה בין פריטים סגורים לפריטים פתוחים

פריטים פתוחים	פריטים סגורים	
ידע, הבנה ויישום, וכן חשיבה מקורית, הבעה, יכולת כתיבה וכדומה	ידע, הבנה, יישום וכדומה	מהן המיומנויות שאפשר לבדוק?
בבניית המחווון ובבדיקת הפריטים	בחיבור הפריטים	עיקר המאמץ
דורשת מחשבה ומומחיות והיא סובייקטיבית	קלה, מהירה, זולה ואובייקטיבית	בדיקת השאלה
בינונית	גבוהה	מהימנות

מהימנות, תוקף והוגנות

מהן תכונותיו של מבחן טוב?

מבחן טוב הוא -

✓מהימן

✓תקף

✓הוגן

מהימנות

מדד לעקביות וליציבות של הציונים:

✓ בין העברות חוזרות

✓ בין מועדים מקבילים

✓ בתנאים משתנים (זמן, מקום, בודק)

✓ בין בודקים שונים

✓ במבחן מהימן, אילו תיעשה המדידה פעם נוספת, יקבל הנבחן ציון דומה.

ככל שהמבחן מהימן יותר, הציון המתקבל משקף טוב יותר את היכולת האמיתית של הנבחן

איך אומדים את המהימנות?

✓ נוסחים מקבילים (Parallel Test reliability)

שני נוסחים מקבילים מועברים לאותה קבוצה באותו המועד או במועדים שונים.

✓ מהימנות מבחן חוזר (Test-retest reliability)

אותו נוסח מועבר לאותה קבוצת נבחנים בשתי הזדמנויות שונות.

✓ מהימנות מבחן חצוי (Split half reliability)

מתאם בין שני חלקי מבחן, למשל פריטים זוגיים ואי-זוגיים.

✓ מתאם בין מעריכים (Inter-rater reliability)

מתאם בין הערכותיהם הבלתי-תלויות של בודקים שונים עבור אותו נבחן.

מהימנות

✓ חישוב המהימנות יכול להיעשות באמצעים סטטיסטיים: לרוב חישובי מתאמים בין פריטים שונים בתוך המבחן, בין מועדים שונים, בודקים שונים ועוד.

✓ מדד המהימנות הוא קורלציה, ולכן נע בין 0 לבין 1.

מהימנות

✓ גם כאשר אין משאבים או צורך למדוד מהימנות,
ישנן דרכים בדוקות לשיפור מהימנותם של
מבחנים:

✓ הגדלת מספר השאלות

✓ שימוש בשאלות בעלות יכולת הבחנה גבוהה

✓ שאיפה למספר עקרונות שנראה בהמשך

מדוע אין מבחן לגמרי מהימן?

✓ גורמי מבחן: ההצלחה בפריטים רבים מקרית, המבחן קצר מדי וכו'.

✓ גורמי העברה: הטמפרטורה בחדר, השעה, העונה.

✓ גורמי נבחן: הנבחן עושה את המבחן ברמות שונות של ערנות, קשב, חרדה, מוטיבציה וכו'.

✓ גורמי בודק: הבודק בודק מבחנים שונים ברמות שונות של ערנות, קשב ומצב רוח; הבדיקה אינטואיטיבית ולא עקבית.

הטיות נפוצות בבדיקת מבחנים

Halo Effect ✓ – חוסר הבחנה בין אספקטים שונים בתפקוד של הנבחן

Primacy Effect ✓ – מידע שמתקבל מוקדם יותר משפיע יותר

Contrast Effect ✓ – הנטייה להשוות כל הערכה לקודמתה

Confirmation Bias ✓ – הנטייה לפרש ולייצר אינפורמציה שמאששת אמונות קדומות

קפדנות יתר או נדיבות יתר – נטייה אישית שונה של בודקים שונים

דעות קדומות וחשיבה סטראוטיפית

תוקף

באיזו מידה המבחן מודד את מה שהוא אמור למדוד.

סוגים של תוקף:

✓ **תוקף תוכן** – המידה שהמבחן מייצג את עולם התוכן שהוא נועד למדוד.

✓ **תוקף תלוי-קריטריון** – המידה שהציון במבחן מתואם עם קריטריון חיצוני רלוונטי.

✓ **תוקף נראה** – המידה שנראה לנבחנים שהמבחן מודד את מה שהוא אמור למדוד.

✓ **תוקף השלכתי** – האופן שהמבחן משפיע על הסביבה, על הלמידה ועל החברה.

חישוב התוקף

✓ תוקף הקריטריון מחושב בדרך כלל לפי המתאם בין
הציון בבחינה לבין הציון בקריטריון.

הקשר בין המהימנות לבין התוקף

✓ יתכן כי מבחן יהיה מהימן אך לא יהיה תקף.

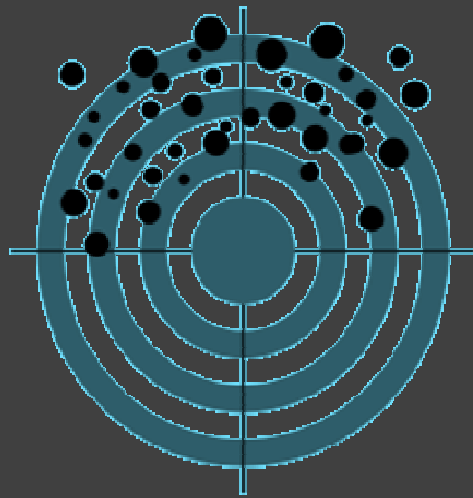
▪ למשל, אם הוא מודד משהו עקבי אך לא את מה שאנחנו חושבים שהוא מודד.

✓ לא יתכן מבחן תקף שאינו מהימן.

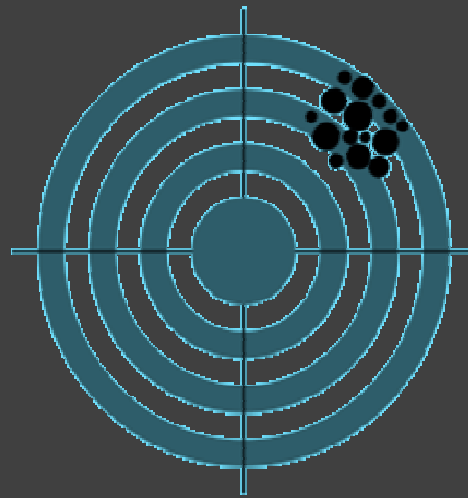
▪ מהימנות היא תנאי הכרחי לתוקף.

✓ מהימנות היא הגבול העליון שיכול להיות לתוקף.

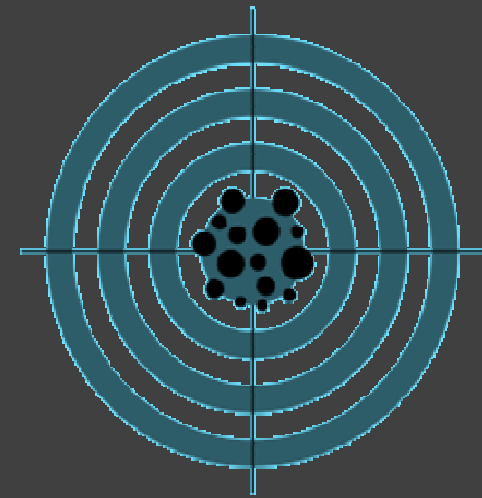
הקשר בין המהימנות לבין התוקף



לא מהימן
ולא תוקף



מהימן ולא
תוקף



מהימן
ותוקף

ממוצע הציונים בבחינה הפסיכומטרית (2011)

משתנה	קבוצה א'	קבוצה ב'	פער
שפת בחינה	ערבית – 466	עברית – 566	100
מצב חברתי-כלכלי	נמוך – 483	גבוה – 591	108
מגדר	נשים – 518	גברים – 561	43
מועד בחינה	אפריל – 525	פברואר – 655	30
גיל	30 ומעלה – 472	בין 20 ל-22 – 552	80

הוגנות

- ✓ במבחן הוגן הבדלים בתכונה הנמדדת יוצרים את ההבדלים בציונים של הנבחנים.
- ✓ מבחן הוגן אינו מפלה בין קבוצות באוכלוסיה, הוא מבטיח שיוויון בין שווים.
- ✓ מבחן הוגן מאפשר לנבחן להפגין את הידע שלו.

סוגיות בהוגנות מבחנים אקדמיים:

- פקטורים
- השוואה בין מועדים
- רגישות לקבוצות מיעוט
- התאמות בבחינות לסטודנטים עם לקויות

סיכום

✓ ככל שהמבחן מהימן יותר ותקף יותר, כך הוא הוגן יותר

✓ יש דרכים בדוקות לשיפור המהימנות והתוקף של
מבחנים