

אפי | جاس | ISPA

אגודה ישראלית לפסיכומטריקה  
الجمعية الإسرائيلية للبيكوميترية  
ISRAELI PSYCHOMETRIC ASSOCIATION  
מבחנים ■ מדידה ■ הערכה

הכינוס ה-18  
של  
האגודה הישראלית  
לפסיכומטריקה

יום ב' ט"ו בשבט תשפ"ב 17.1.2022



כנס מקוון



2022

בתמיכת:



מרכז ארצי לבחינות ולהערכה (נ"ר)  
NATIONAL INSTITUTE FOR TESTING & EVALUATION  
المركز القطري للامتحانات والتقييم  
מיסודן של האוניברסיטאות בישראל

האגודה הישראלית לפסיכומטריקה (אפי) היא גוף המאגד בתוכו אנשי מקצוע מתחום המדידה וההערכה החינוכית והפסיכולוגית.

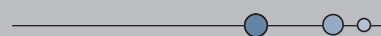
מטרות האגודה:

- לעודד את הפעילות בתחום הפיתוח והיישום של מבחנים סטנדרטיים ושיטות להערכת ביצוע
- לקדם את המחקר והעשייה בתחומי המדידה החינוכית והפסיכולוגית
- לשכלל את שיטות ההערכה וכלי המדידה ולבסס את יישומם המושכל
- להתוות, להעריך ולהפיץ כלים מתודולוגיים עדכניים ושיטות הערכה חדשות
- להתריע על שימוש לקוי במבחנים ועל פרשנות מוטעית של תוצאותיהם
- לעודד קשר מקצועי, חילופי מידע ושיתוף פעולה בין העוסקים בתחום המדידה וההערכה

# אפי



<p><b>ברכות ופתיחה</b> 8:55 – 9:20</p> <p>דברי פתיחה – יו"ר הכנס: גילי גולדצויג, חנן גולדשמידט                  פתיחה יו"ר אפי היוצא: צור קרליץ                  יו"ר אפי הנכנסת: ריטה סבר</p>	
<p><b>הרצאת אורח</b> 9:20 – 10:20</p> <p>האתגר הפסיכומטרי במדידת יצירתיות                  יועד קנת</p>	
<p><b>הפסקה</b> 10:20 – 10:25</p>	
<p><b>מושב א'</b> 10:25 – 11:15</p> <p>פסיכומטריקה וסטטיסטיקה                  יו"ר: נעמי גפני</p> <ul style="list-style-type: none"> <li>■ אמידת פרמטרים לפריטי האנגלית שהועברו במאל"ו בשנים 1990–2020 מיכל באומר, לאה בורובוי, יואל רפ</li> <li>■ בדיקת תלויות בתוך מקבץ פריטים בגישת IRT צור קרליץ, נעם כאהן</li> </ul>	
<p><b>הענקת פרס אפי והצגת העבודה הזוכה</b> 11:15 – 11:40</p> <p>יו"ר ועדת הפרס: רוני ברצלבסקי</p>	
<p><b>הפסקה</b> 11:40 – 11:45</p>	
<p><b>מושב ב'</b> 11:45 – 12:35</p> <p>אמידה וקריטריונים                  יו"ר: רות פורטס</p> <ul style="list-style-type: none"> <li>■ מדידה של חשיבה יצירתית באמצעות מטלות מחשב אינטראקטיביות מירי ברחק-רבינוביץ, צור קרליץ, עדו רול</li> <li>■ אמידת תוקף ניבוי באוכלוסייה לפני הליך קבלה ונשירה דביר קלפר, איתמר כהן, תמר קנת-כהן, צור קרליץ</li> </ul>	
<p><b>הפסקת צהריים – פיצול לחדרים למפגשים חברתיים</b> 12:35 – 13:15</p>	
<p><b>מושב ג'</b> 13:15 – 14:05</p> <p>תהליכי קריירה                  יו"ר: חגי רבינוביץ'</p> <ul style="list-style-type: none"> <li>■ תוקף המבנה של שאלון הקשיים בבחירת תחום לימודים ומקצוע: היבטים מתודולוגיים בניתוח מעבר לעשר שפות נמרוד ליון, שגיני אודיאר, יוליה ליפשיץ-ברזילר, איתמר גתי, ג'רום רוסייר</li> <li>■ מה מסביר טוב יותר את הפער המגדרי בבחירת לימודים מדעיים? מאיר ברנרן, יואל רפ</li> </ul>	
<p><b>הפסקה</b> 14:05 – 14:10</p>	
<p><b>פאנל מומחים במדידה והערכה</b> 14:10 – 15:30</p> <p>מארגן ומנחה: אבי אללוף                  משתתפים: עינב אבני, ענת בן סימון, גרשון בן שחר, אייל גמליאל, רענן הס</p>	
<p><b>אסיפת חברי האגודה</b> 15:30</p>	



## האתגר הפסיכומטרי במדידת יצירתיות

יועד קנת

הפקולטה להנדסת תעשייה וניהול, הטכניון

יצירתיות הינה תכונה אנושית מוערכת ונחשקת. למרות זאת, שנים רבות יצירתיות לא נחקרה באופן מדעי מהותי. הסיבות לעיסוק המדעי המועט ביצירתיות הינן קושי בהגדרה של תכונה זו ובמיוחד בכלי מדידה והערכה שלה. למעשה, כלי המדידה והערכה של תכונה זו במעבדה מבוססים באופן כמעט מוחלט על שיפוטיות סובייקטיביים: שופטים מתבקשים לשפוט עד כמה תשובות במטלה יצירתית הינן "יצירתיות", תהליך זה מייצר אתגרים ושאלות לגבי התקפות ומהימנות של טיב השיפוט. בשני העשורים האחרונים, חלה עלייה משמעותית בחקר המדעי של יצירתיות, בשל התמקדות במחקר קוגניטיבי ומוחי של תופעה מורכבת זו. התמקדות זו הובילה לשיפור ופיתוח של נושא מדידת והערכת יכולת יצירתית. במיוחד, הצמיחה בפופולריות של שימוש בכלים מבלשנות חישובית ומודלים מתמטיים לחקר יצירתיות הולך ומוביל את התחום לשימוש בכלי מדידה כמותיים ואובייקטיביים בהערכת הבדלים בין אישיים ביכולת יצירתית. בהרצאה זו אציג כיצד השימוש בשיטות חישוביות אלו מאפשרות לחקור את התפקיד של ידע בחשיבה יצירתית. זאת, באמצעות הצגת מחקרים המראים כיצד לאנשים יותר יצירתיים יש מערכת ידע יותר גמישה המאפשרת להם לחפש באופן יותר יצירתי במערכת זו. אסיים בדיון באתגרים והזדמנויות בשימוש בכלים אלו בחקר יצירתיות בעברית.

## אמידת פרמטרים לפריטי האנגלית שהועברו במאל"ו בשנים 1990 – 2020

מיכל באומר, לאה בורובוי, יואל רפ  
מאל"ו

לאחרונה שוקד המרכז הארצי לבחינות ולהערכה על פיתוח מבחן חדש באנגלית כשפה זרה. במבחן זה יבדקו המיומנויות קריאה, הבנת הנשמע וכתיבה, הוא צפוי להיות אדפטיבי ויועבר באמצעות מחשב למועמדים להשכלה גבוהה באופן מתמשך לאורך כל השנה. על מנת לבנות מבחן זה, התעורר הצורך לאפיין את כל פריטי האנגלית אשר הועברו במשך התקופה בת 30 השנים במבחנים קודמים. אפיון זה יאפשר להתאים טוב יותר את פריטי המבחן לאוכלוסייה מגוונת ויאפשר חקירה היסטורית של התפלגויות מאפייני הפריטים ושל התפלגויות הידע באנגלית באוכלוסייה בישראל.

לאור העובדה כי הפריטים הנאמדים נכתבו והועברו לאורך תקופה ארוכה ביותר, מצבור הפריטים היה עצום וכלל כמות של מעל 12 אלף פריטים שהועברו למעל מיליון שלוש מאות אלף נבחנים. תוכנות מדף סטטיסטיות אינן יכולות להתמודד עם כמות נתונים גדולה כל כך לאמידת פרמטרים. אי לכך, היה צורך לפתח שיטה ייחודית לאמידת הפרמטרים של מאגר פריטים אדיר ממדים זה.

הפריטים: 12,096 פריטים שהועברו בחלק האנגלית של הבחינה הפסיכומטרית ובבחינת אמי"ר בין השנים 1990 – 2020.

חלקנו את מצבור הפריטים הזה ל-14 קבוצות מקריות בדגימת שכבות תוך שמירה על הפרקים שבהם הם נכללו. כך נוצרו קבוצות עם ייצוג פרופורציונאלי של סוג הבחינה (פסיכומטרי ואמי"ר), שנת העברה ושפת הבחינה (עברית וערבית). בכל קבוצת ניתוח היו כ-1,000 פריטים, שהשיבו עליהם כ-300,000 נבחנים. בכל קבוצה הוספנו שני "פרקי עוגן" (כ-50 פריטים), שחזרו ונכללו בכל אחת מ-14 קבוצות ניתוח. כמו כן כדי לאשש את הליך אמידת הפרמטרים ולעמוד על מידת היציבות של ההליך, ערכנו שתי רפליקציות מלאות להליך זה. בתוך כל קבוצה, אמדנו את שלושת הפרמטרים של הפריטים (הבחנה, קושי ורמת ניחוש) במודל IRT התלת-פרמטרי (PL3) ואת רמות היכולת של הנבחנים (ציוני התטה). ניתוח התוצאות מלמד על תקפות ההליך ועל מידת יציבות טובה ומציג עדויות אמפיריות לתקינותו הכוללת של ההליך.

## בדיקת תלויות בתוך מקבץ פריטים בגישת IRT

צור קרליץ, נעם כאהן  
מאל"ו

במבחנים רבים נהוג להשתמש במקבצי פריטים המתייחסים לאותו גריין. לדוגמה, מספר פריטי הבנת הנקרא המבוססים על טקסט אחד (מקבץ כזה נקרא בשם טסטלט, Testlet). במצב זה נצפה שתהיה תלות בין הביצוע בכל הפריטים המתייחסים לאותו הגריין. לדוגמה, עולה חדש שיקבל טקסט המבוסס על התרבות הפופולרית בארץ, אותה הוא עשוי כלל שלא להכיר, עלול להיתקל בקושי כללי עם הטקסט, אשר ישפיע על הסיכוי שלו לענות נכונה על כל הפריטים המוסבים על אותו הטקסט. במילים אחרות, חלק מהשונות בביצוע בפריטים אלה מוסברת על ידי גורם שאינו היכולת הנמדדת (במקרה זה הכרות עם התרבות בארץ, אם כי בד"כ מדובר במאפייני הטקסט – אורך, איכות, משלב). אם לא נתחשב בגורם הנוסף הזה כשננתח את נתוני המבחן אנו עלולים לקבל אומדנים מוטים של המהימנות, קושי הפריטים ויכולת הנבחנים (Wainer & Lukhele, 1997). על כן, חשוב לבחון את מידת השפעתו של הגורם הנוסף, וזאת על ידי בדיקת קיומה של תלות בין הפריטים באותו המקבץ.

המחקר הנוכחי מתבסס על תורת התגובה לפריט (IRT), לבדיקת מידת התלות הפנימית במקבצי פריטים מבוססי גריין משותף. אחת ההנחות היסודיות, החזקות והמשמעותיות ביותר של IRT היא הנחת האי-תלות המקומית. משמעותה של הנחה זו היא כי בהינתן יכולת קבועה, קיימת אי-תלות בין התגובות לפריטים השונים. במילים אחרות, הדבר היחיד שמשפיע על הסיכוי לענות נכון על הפריט היא יכולת הנבחן ואין גורם אחר שמשפיע על הביצוע. אם היה גורם נוסף, היינו רואים תלות בין הביצוע בפריטים שונים, גם לאחר שהחזקנו את יכולת הנבחן קבועה. הנחת האי-תלות המקומית היא הכרחית לקבלת אומדנים לא מוטים של הפרמטרים במודל IRT ולכן יש לבדוק שהיא מתקיימת בכל ניתוח מסוג זה. סביר לצפות שהנחת האי-תלות המקומית תופר במצבים לעיל, בהם ישנם מספר פריטים המבוססים על גריין אחד שעשוי ליצור תלות בין הפריטים באופן שלא קשור ליכולת הנמדדת.

במחקר הנוכחי השתמשנו בשתי שיטות מבוססות IRT לבדיקת תלויות בתוך מקבצי פריטים:

**חישוב מדד Q3** של Yen (1984). מדד זה בוחן את המתאם המנוכה בין זוגות פריטים, כלומר את המתאם מעבר לקשר המוסבר על ידי היכולת הנמדדת. מתאם מנוכה השונה מ-0 באופן מובהק מצביע על כך שבין הפריטים ישנו קשר שאינו מוסבר על ידי יכולת הנבחנים, או לענייננו – הנחת האי-תלות המקומית מופרת.

**אמידת מודל טסטלט** (Testlet Model – Wang, Bradlow & Wainer, 2002). המודל אומד, בנוסף לקושי הפריט, גם פרמטר המייצג את האינטראקציה בין התלמיד לגריין. בדיקת הנחת האי-תלות מבוססת על בחינת שונות הפרמטרים של האינטראקציה בכל טסטלט.

המחקר ממחיש את השימוש בשתי השיטות לעיל בעזרת נתוני אמת ממבחני המיצ"ב ומהבחינה הפסיכומטרית. ההרצאה נועדה לסייע בהעלאת המודעות לבעיות שעלולות להיווצר במבחנים מהסוג הזה ובהפצת הידע לגבי הדרכים לבדוק ולטפל בבעיות אלה.

## מדידה של חשיבה יצירתית באמצעות מטלות מחשב אינטראקטיביות

מירי ברחק-רבינוביץ<sup>1</sup>, צור קרליץ<sup>2</sup>, עדורול<sup>1</sup>

<sup>1</sup> הטכניון

<sup>2</sup> מאל"ו

דמות הבוגר של משרד החינוך למערכת החינוך של שנת 2030 שמה למטרה לפתח מיומנויות מורכבות בקרב תלמידי בית הספר כגון הכוונה עצמית, חשיבה ביקורתית, חשיבה יצירתית ועוד. על מנת לקדם מיומנויות אלה בקרב תלמידי בית הספר, עולה הצורך לפתח מתודות מדידה חדשות, אשר יאפשרו מדידה של יכולות (נוסף על ידע). אתגר מדידה זה נמצא בבסיס של מבחני פיזה בנושאים המיוחדים, המתמקדים בבחינה הקרובה במדידה של חשיבה יצירתית כמיומנות מורכבת.

חשיבה יצירתית הוגדרה במבחני פיזה 2022 כיכולת לייצר, להעריך ולשפר פתרונות על מנת להגיע לתוצאות טובות יותר בבעיה קשה לפתרון. זהו תהליך איטרטיבי ודינמי של חשיבה, שבו הנבדק מתנסה שוב ושוב ביצירה מחדש של פתרונות ושיפור פתרונות קודמים. בכך ההגדרה שמה את עיקר הדגש על התהליך של חשיבה יצירתית על פני התוצר.

הגדרה זו של פיזה הולמת את הגדרתו של Torrance, המוכר כאבי היצירתיות, אשר פיתח באמצע המאה הקודמת את המבחנים הקלאסיים למדידה של חשיבה יצירתית. מבחנים אילו עדיין נמצאים בשימוש נרחב באקדמיה, בחינוך ובמערכות המיון. כך למשל, נבדק מתבקש לצייר כמה שיותר פריטים עם משמעות מעיגול ריק או לציין מספר רב של שימושים ללבנה. עם זאת, נשאלת השאלה האם מבחנים אילו אכן מודדים חשיבה יצירתית כמיומנות מורכבת, וכיצד ניתן להעריך את התהליך האיטרטיבי של חשיבה יצירתית.

מטלות מחשב אינטראקטיביות מאפשרות מספר חידושים במדידה המאפשרים להעריך מיומנויות מורכבות כגון חשיבה יצירתית. תחום אחד של חדשנות הינו אופי המטלה עצמה. מטלות ממוחשבות יכולות לכלול עזרים וכלים שונים, וכן לכלול משוב בזמן אמת לנבדקים. באופן זה יש הזדמנות לנבדקים ליישם תהליך איטרטיבי ודינמי של חשיבה, כך שנבדקים יכולים להעריך ולשפר פתרונות קודמים. תחום החדשנות השני הינו שימוש במידע תהליכי (process data). מטלות מחשב אינטראקטיביות מתעדות כל פעולה שהנבדק עשה במערכת. תעוד זה יכול לשמש אותנו להערכת תהליך החשיבה שביצע הנבדק. ואולם, על מנת להפיק משמעות מנתונים אלו יש צורך בהבנייה סדורה של פרקטיקות היודעות לנתחם. תחום המחקר בשם Learning analytics מאגד טכניקות שיוודעות לפעול על מידע תהליכי, בין היתר לצרכי הערכה ומדידה.

במסגרת המחקר הנוכחי ביצענו מעבדות קוגניטיביות עם 9 תלמידים, במהלכן בדקנו 5 מטלות שונות להערכת חשיבה יצירתית. לדוגמא, מטלה אחת ביקשה מתלמידים לבנות פנס חזק בהינתן רכיבים מוגדרים, מטלה אחרת הנחתה אותם לפתח נוסחה למדידת עקביות, ומטלה שלישית ביקשה מהם לעצב גלגל תחת מערכת של אילוצים. תהליך העבודה של התלמידים נותח על מנת לענות על שאלות המחקר הבאות: (1) מהן הדרישות ממטלות אינטראקטיביות בהן תלמידים יכולים להביא לידי ביטוי את מגוון האספקטים של חשיבה יצירתית? (2) כיצד מתקפים מבחן מחשב אינטראקטיבי? (3) כיצד מצייננים אוטומטית מיומנות מורכבת? ההרצאה תסקור את הממצאים מהניסוי. על סמך התנסות זו, נגדיר את האפשרויות והאתגרים במדידה של מיומנות מורכבת באמצעות מטלות מחשב אינטראקטיביות, בהסתמך על מקרה בוחן של מדידת חשיבה יצירתית.

## אמידת תוקף ניבוי באוכלוסייה לפני הליך קבלה ונשירה – השוואה בין שלוש שיטות לטיפול בנתונים החסרים

דביר קלפר, איתמר כהן, תמר קנת-כהן, צור קרליץ  
מאל"ו

במחקרי תוקף ניבוי משתמשים לרוב בנוסחת התיקון לקיצוץ תחום על מנת לנטרל את ההשפעה של הליך הקבלה על המתאם בין החזאי לקריטריון. נוסחה זו מתאימה למצב שבו קיימת ברירה של תצפיות לפי החזאי, כאשר הנמוכים בחזאי אינם מתקבלים ללימודים ולפיכך אין להם ציון בקריטריון. במקרה שקיימת גם נשירה, כלומר שציוני קריטריון חסרים שלא בגלל ציון בחזאי, הנחות הנוסחה לתיקון לקיצוץ תחום לא בהכרח מתקיימות.

מחקר זה בדק בעזרת סימולציות את איכות אמידת המתאם בין החזאי לקריטריון במצב שבו בנוסף להליך הקבלה קיימת גם נשירה. המחקר השווה בין שלוש שיטות לטיפול בנתונים החסרים: נוסחת התיקון לקיצוץ תחום, שיטת ההשלמות המרובות, שמיועדת לטפל במצב של נתונים שחסרים באופן אקראי (MAR), ומודל הברירה של הקמן, שמיועד לטפל במצב של נתונים שחסרים שלא באופן אקראי (MNAR). המחקר יישם שלוש דרכים למידול נשירה: נשירה אקראית, נשירה ודאית של החלשים בקריטריון, ומצב ביניים של "נשירה מתואמת קריטריון" (מתאם שלילי, אך לא מושלם, בין נשירה לקריטריון). איכות האמידה נבדקה תוך שליטה בפרמטרים כגון גודל האוכלוסייה, המתאם בין החזאי לקריטריון באוכלוסייה, שיעור הנדחים ושיעור הנושרים. בנוסף להצלבות המלאות של כל הפרמטרים לעיל, נבדקה איכות האמידה כאשר יש משתנה עזר – משתנה שמתואם עם הקריטריון או עם הנטייה שיהיה נתון חסר בקריטריון. הגדרנו שתי חלופות למשתנה העזר: משתנה מסוג "יכולת התמדה" שמתואם יותר עם הקריטריון מאשר עם החזאי, ומשתנה מסוג "מהירות" שמתואם יותר עם החזאי מאשר עם הקריטריון.

התוצאות מראות שמעבר לתנאים השונים, התיקון לקיצוץ תחום הינה השיטה העדיפה, במובן שבמרבית התנאים היא מדייקת, וכשלא – היא מספקת אומדן שמרני של המתאם. שיטה זו מדייקת פחות ככל שיש יותר נשירה שאינה אקראית. שיטת ההשלמות המרובות נותנת תוצאות דומות לאלו של התיקון לקיצוץ תחום, ובנסיבות מסוימות (בפרט, כשמשתמשים במשתנה עזר מסוג "יכולת התמדה") – היא מדייקת יותר, אך במחיר של מורכבות חישובית. מודל הברירה של הקמן עדיף רק במקרים של אוכלוסייה גדולה, עם נשירה ודאית של החלשים ועם אחוז גבוה יחסית של נשירה לעומת דחייה, אם כי השימוש בו עשוי להיות כרוך באומדן יתר של המתאם.

המחקר כלל גם שני ניתוחים משניים. בראשון, בדקנו את התלות של איכות האמידה במשתנה העזר באופן יותר שיטתי ומקיף. בבדיקה זו נמצא, שכצפוי, המתאם של משתנה העזר עם הקריטריון משפיע על איכות האמידה יותר מאשר המתאם שלו עם החזאי. בניתוח המשני השני, השווינו את הפיזור באומדני שתי שיטות – תיקון לקיצוץ תחום והשלמות מרובות – במצב שבו שתי השיטות אומדות את המתאם באופן כמעט מושלם (נשירה אקראית). בדיקת הפיזור מעבר לסימולציות השונות גילתה יתרון לשיטת ההשלמות המרובות עם משתנה עזר מסוג "יכולת התמדה".



## תוקף המבנה של שאלון הקשיים בבחירת תחום לימודים ומקצוע: היבטים מתודולוגיים בניתוח מעבר לעשר שפות

נמרוד לוי<sup>1,2</sup>, שגיני אודיאר<sup>2</sup>, יוליה ליפשיץ-ברזילר<sup>1</sup>, איתמר גתי<sup>1</sup>, ג'רום רוסייר<sup>2</sup>

<sup>1</sup> האוניברסיטה העברית בירושלים, ישראל

<sup>2</sup> אוניברסיטת לוזאן, שוויץ

בחירת תוקף מבנה הוא אחד המבחנים האמפיריים הראשונים בתהליך פיתוח של כלי מדידה. ניתוח גורמים מאשש (CFA) מהווה את השיטה האנליטית המקובלת לבחינת המידה בה מערכת הקשרים בין פריטי הכלי תואמת את המבנה המשוער של המבנה הנמדד. Hopwood & Donnellan (2010) מציינים כי בספרות ניתן למצוא דוגמאות רבות לבעיות ביישום של ניתוח גורמים. בהרצאה זו נדון בכמה מהם באמצעות הדוגמא של בחירת תוקף המבנה של שאלון הקשיים בבחירת תחום לימודים ומקצועי.

שאלון הקשיים בבחירת תחום לימודים ומקצוע פותח לזיהוי מוקדי הקושי של מתלבטים בבחירת תחום לימודים ומקצוע (Gati, et al., 1996). בשאלון 32 פריטים המייצגים 10 קטגוריות קושי, המקובצות בשלושה אשכולות ומאפשרים גם חישוב ציון כללי (1-3-10-32). בתהליך פיתוח השאלון ניתוח גורמים מגשש (EFA) תמך במבנה הרב-ממדי המשוער של השאלון, אולם פיתוח השאלון נערך טרם החל השימוש הנפוץ בניתוח גורמים מאשש (CFA). בחלק מהמחקרים שנערכו בהמשך לתיקוף של מבנה השאלון בשפות שונות נמצאו סטיות מהמבנה התיאורטי המקורי. כך למשל, נתונים שנאספו באיראן, בדרום קוריאה, בסין, ובקרוואטיה הובילו להעדפה של מבנה תיאורטי שאינו כולל את כל הממדים. במקרים אחרים – לרבות בישראל, בשוויץ, ובתורכיה – תיקוף המבנה התיאורטי המקורי של שאלון הקשיים נעשה ברמת הסולמות המייצגים את 10 קטגוריות הקושי תוך התעלמות מרמת הפריטים.

במחקר שערכנו לבחינה מחודשת של תוקף המבנה של שאלון הקשיים, ניתחנו נתונים שנאספו במסגרת 52 מחקרים (21,594 נבדקים) בעשר שפות מ-19 מדינות. לצורך בחינת מבנה השאלון, נבדקו והשווה ארבעה מודלים מבניים: (1) המודל המקורי הכולל עשרה סולמות, שלושה אשכולות, וציון כללי, (2) מודל הכולל רק עשרה גורמים וציון כללי, (3) מודל הכולל רק שלושה אשכולות וציון כללי, (4) ומודל הכולל רק ציון כללי. ניתוח מקדים של מאפייני הפריטים הצביע על סטייה משמעותית של התפלגות פריטי השאלון מנורמליות. על כן, השימוש באמדתן Maximum Likelihood (ML), ברירת המחדל במרבית התוכנות לניתוח גורמים המצריך את הנחת הנורמליות של הפריטים, אינו מתאים. במקומו, השווינו את ארבעת המודלים באמצעות אמדתן מסוג Weighted Least Squares with Mean and Variance Adjusted (WLSMV) שאינו מניח התפלגות נורמלית של הפריטים. בעקבות השימוש באמדתן זה, תוקף המבנה המקורי של השאלון אושש (התאמתו של מודל 1 לנתונים נמצאה טובה יותר משל המודלים החלופיים), והמהימנות של סולמות השאלון נמצאה בטווח הרצוי.

ניתוח המבנה של שאלון הקשיים בבחירת תחום לימודים ומקצוע בעשר שפות אישש את המבנה התיאורטי המקורי שהציעו Gati et al. (1996), וזאת בניגוד למחקרי התוקף הקודמים. ברמת המבנה הנמדד, ממצאים אלה תומכים במסקנה שניתן להבחין בין עשרה מוקדי קושי המקובצים בשלושה אשכולות ומאפשרים חישוב ציון כולל. ברמה המתודולוגית, ממצאי המחקר מדגישים את החשיבות של (1) בחינת מאפייני הפריטים בטרם ביצוע ניתוח גורמים מאשש על מנת לוודא שימוש באמדתן המתאים, (2) שימוש במדגמים הצפויים לכלול שונות מספקת במשתנה הנמדד, ו-(3) ניתוח ברמת הפריטים (ולא הסולמות) על מנת לוודא בחינה מלאה של המאפיינים הפסיכומטריים של הכלי הנבחן.

## מה מסביר טוב יותר את הפער המגדרי בבחירת לימודים מדעיים?

מאיר ברנרן, יואל רפ  
מאל"ו

במדינות המערב, נשים זוכות לייצוג נמוך בלימודים גבוהים ובתעסוקה בתחומי המדע, הטכנולוגיה, ההנדסה והמתמטיקה (STEM). תופעה זו מהווה מקור לדאגה בשל המחסור באנשי מקצוע מיומנים בתחומים אלו. בנוסף, מכיוון שתחומים אלו מתוגמלים היטב לעומת תחומי תעסוקה אחרים, ייצוג נשי נמוך במקצועות STEM תורם לפערי השכר המגדריים.

בעבר, חוקרים סברו כי מקורה של תופעה זו ביתרונם של בנים בהישגים הלימודיים בתחום המתמטיקה בפרט ובחשיבה כמותית בכלל. בשנים האחרונות מסתמנת בעולם מגמה של צמצום עד סגירה של ממש של פערים אלו אך הדבר לא בא לידי ביטוי בסגירת פערי הייצוג במקצועות ה-STEM.

לאחרונה, בעקבות הצטברות נתונים ממחקרים בינלאומיים חינוכיים, נמצא כי בנים נוטים להיות טובים יותר במתמטיקה מאשר בקריאה, בעוד שבנות טובות יותר בקריאה מאשר במתמטיקה. כתוצאה מהיתרון היחסי של בנות בקריאה (ומקצועות עתירי מלל באופן כללי) והיתרון היחסי של בנים במתמטיקה (ומקצועות עתירי מתמטיקה בכלל), חוקרים העלו את ההשערה כי נשים רואות את עצמן טובות יותר בחשיבה מילולית ופחות בחשיבה כמותית. מסיבה זו, בנות יעדיפו פחות ללמוד ולעסוק בתחומי ה-STEM.

המחקר הנוכחי מאתגר השערה זו. במסגרת המחקר נאספו נתונים של כ-620,000 נבחנים בבחינה הפסיכומטרית בישראל ב-30 שנים האחרונות, ואת בחירותיהם ההשכלתיות. נמצא כי היתרון היחסי של גברים בחשיבה כמותית (ושל נשים בחשיבה מילולית) אכן תורם להגדלת הפער המגדרי בבחירת לימודי STEM. עם זאת, נמצא כי ההישגים בחשיבה כמותית כשלעצמם, ולא רק ביחס להישגים בחשיבה מילולית, משפיעים על הפער הזה כמעט באותה מידה. המסקנה שלנו היא, אפוא, שדי להתמקד בהסבר הפשוט יותר, דהיינו הבדלים בין מגדריים בחשיבה כמותית, כאשר מבקשים לעמוד על המקורות לפערי הייצוג המגדריים במקצועות ה-STEM.

## השפעת המבחנים על החברה ולהפך

מארגן ומנחה: אבי אללוף - מאל"ו

משתתפים:

עינב אבני - ענף מיון, צהל ענת בן סימון - מאל"ו, גרשון בן שחר - האוניברסיטה העברית

אייל גמליאל - מרכז אקדמי רופין, רענן הס - עצמאי, פסיכולוג תעסוקתי

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 וקביעת 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 ההסמכה חסרים? האם המבחנים מנציחים פערים בחברה? האם הם הוגנים? כיצד ייראו המבחנים בעתיד? האם נפוצות ההכנה למבחנים פוגעת בתהליכי מיון? האם חשוב לחזק את האוריינות בציבור בנושאי מדידה והערכה? 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.

פאנל של חמישה משתתפים בעלי מומחיות במדידה והערכה, הבאים מרקע שונה, מהאקדמיה, מארגונים ציבוריים וחברות פרטיות, ישיבו על שאלות אחידות, מתוכננות מראש. השאלות יסוכמו מראש עם המשתתפים מתוך מטרה לאפשר הצגת גישות שונות ומגוון דעות והשקפות מבוססות נתונים. יוקצה זמן מוקצב לתשובה של כל אחד מחברי הפאנל.

להלן דוגמאות לשאלות אפשריות בפאנל (1) מה דעתך על שימוש ב"טביעת האצבע הדיגיטלית" (ובכלל זה פייסבוק, לינקדין וכו') למטרות אבחון, מיון וקבלה? (2) מה דעתך על תופעת ההכנה לבחינות? (3) מהי ההשפעה שיש למבחנים על הכלכלה? (4) האם תוכלו להצביע על מקרים שבהם המבחנים גרמו נזק לחברה?

לאחר שחברי הפאנל יתייחסו לשאלות שיופנו אליהם יתקיים דיון. הקהל יוזמן להתייחס לדברים שנאמרו ולהפנות שאלות לחברי הפאנל.

## אנשי קשר ומציגים בכינוס אפי 2022

דוא"ל	שם	
tzur@nite.org.il	צור קרליץ	יו"ר אפי
giligold@mta.ac.il	גיל גולדצויג	יו"ר הכינוס
cgoldsch@idc.ac.il	חנן גולדשמידט	
yoedk@technion.ac.il	יועד קנת	הרצאת אורח
gafni.naomi@gmail.com	יו"ר - נעמי גפני	מושב א' פסיכומטריקה וסטטיסטיקה
Leah@nite.org.il	לאה בורובוי	
michalb@nite.org.il	מיכל באומר	
yoel@nite.org.il	יואל רפ	
noam@nite.org.il	נעם כאהן	
fortus@nite.org.i	יו"ר - רות פורטס	מושב ב' אמידה וקריטריונים
Miri.barhak@gmail.com	מירי ברחק	
dvir@nite.org.il	דביר קלפר	
tami@nite.org.il	תמר קנת-כהן	
itamarco@nite.org.il	איתמר כהן	
hagaira@gmail.com	יו"ר - חגי רבינוביץ	מושב ג' תהליכי קריירה
levin.nimrod@gmail.com	נמרוד לוין	
Shagini.Udayar@unil.ch	שגיני אודיאר	
yuliya.lipshits@mail.huji.ac.il	יוליה ליפשץ-ברזילר	
itamar.gati@huji.ac.il	איתמר גתי	
Jerome.Rossier@unil.ch	ג'רום רוסייר	
meir.barneron@gmail.com	מאיר ברנרון	
avi@nite.org.il	מנחה אבי אללופ	פאנל מומחים
eynavavni@gmail.com	עינב אבני	
anat@nite.org.il	ענת בן סימון	
mskpugb@mscc.huji.ac.il	גרשון בן שחר	
eyalg@ruppin.ac.il	אייל גמליאל	
raanan@checkme.co.il	רענן הס	
ispaisrael@gmail.com	מייל לפניות למועצת אפי	
metal@nite.org.il	מיטל אילון	

תגובת והצעת  
קנושא האגודה הישראלית לפסיכומטריקה

נא להפנות אלי:

ד"ר צור קרליץ, יו"ר אפי [ispaisrael@gmail.com](mailto:ispaisrael@gmail.com)



אתר אפי [www.ispa.org.il](http://www.ispa.org.il)



עיצוב ועריכה גרפית | יעל פרץ | מאל"ו