

תמצית

בקרת איכות דו-רמתית על ציונים לפי

משתנים דמוגרפיים

דביר קלפר, אבי אללוף, אליוט טורוול, כרמל אורן, מרינה פרונטון

חישוב ציונים בבחינות רחבות היקף ועתירות סיכון הינו תהליך רב-שלבי מורכב ולכן עשויות לחול בו טעויות שמחירן עלול להיות גבוה לכל הגורמים המעורבים בבחינה. מכיוון שכך, חשוב לבצע פעולות של בקרת איכות קפדנית על ציונים לפני דיווחם לנבחנים ולמוסדות (Allalouf, 2007; Kolen & Brennan, 2014). אחת משיטות הבקרה היא להיעזר בנתונים ידועים מהעבר על קבוצות נבחנים ולהשוות אותם לציוני הקבוצות שהתקבלו בחישוב החדש. במחקרים רבים נמצא קשר בין נתוני רקע דמוגרפיים (מין, גיל, השכלת הורים, הכנסה וכו') להישגים בבחינות שונות (למשל Liu et. al. 2012), כולל הבחינה הפסיכומטרית (סער ואורן 2014). קיומו של קשר זה מאפשר לנבא את ציון הבחינה, הן זה של נבחן בודד והן הממוצע של קבוצת נבחנים וזאת על מנת לאמת את הציונים שהתקבלו בתהליך החישוב. צפוי כי ניבוי פרמטרים קבוצתיים (של הנבחנים בנוסח מסוים במועד בחינה מסוים), כמו ממוצע וחציון, יהיה מדויק יותר (וטעות הניבוי תהיה קטנה יותר), מאשר ניבוי ציון של נבחן בודד. מטרת המחקר הנוכחי היא לבדוק ולהשוות את השימויות של שלושה מודלים רב-רמתיים לניבוי הציון בבחינה הפסיכומטרית על בסיס משתני רקע שונים לצורך בקרת איכות של תהליך חישוב הציונים.

בהתאם לכך בוצע ניתוח היררכי לינארי בשתי רמות כדי לחקור את הקשר של נתוני הרקע לציון הכללי מעבר לנוסחים, כאשר הרמה הראשונה היא הנבחן הבודד, והרמה השנייה היא נוסח הבחינה.

שלושת המודלים שנבדקו הם:

1. ניתוח שונות עם אפקטים מקריים
2. רגרסיה של הממוצעים
3. מודל רגרסיה עם מקדמים מקריים

מודל 1 שימש כבסיס להשוואה עם שני המודלים האחרים. נמצא כי ברמת הנוסח יכולת ניבוי הממוצע היא טובה מאוד (מודל 2 הצליח להסביר 69% מהשונות של ממוצעי הנוסחים). עוד נמצא כי היכולת לנבא ציון של נבחן בודד היא חלשה יחסית (מודל 3 הצליח להסביר רק 20% מהשונות של הציון של נבחן בודד). מודל 2 עם יכולת הניבוי הטובה יותר נבדק על בסיס נתונים שכלל נוסחים שונים מאלה שנכללו במדגם חישוב הפרמטרים ונמצא תקף.

מחקר זה הינו מחקר ראשון המשתמש בניתוח רב-רמתי כדי לחקור את הקשר בין משתני הרקע של נבחנים לבין הציון הכללי בבחינה הפסיכומטרית. תוצאות המחקר מעודדות ונותנות ביד החוקר כלי רב-עוצמה לבקרת איכות של ציונים בבחינה זו כמו גם במגוון רחב של מבחנים שבהם יש בידי החוקר נתונים דמוגרפיים של הנבחנים. בקרת האיכות יכולה להתבצע הן לאחר דיווח הציונים במטרה לבקר את איכות תהליך חישוב הציונים והכילול, והן לפני דיווח הציונים כדי לאתר בעיות בזמן אמת.