

פוסט 4: הדגמה לדחיית תוצאות של מאמר על הטיה של סקרי הוראה, על סמך רשימת הפגמים בפוסטים 1-2

פוסט זה נכתב כתגובה למייל ששלח אחד המשתתפים ברשימה זו ובו הציג מאמר הטוען להטיה מגדרית של סקרי ההוראה:

Mengel, F. and Ulf Zolitz, J.S. (2019). Gender bias in teaching evaluations. *Journal of the European Economic Association*, 17 (2), 535–566. <https://academic.oup.com/jeea/article/17/2/535/4850534>

ציטוטים מדברי הקורא: "אם המחקר מראה שיש הטיה מגדרית בסקרי ההוראה - שנשים מקבלות באופן עקבי ציונים נמוכים יותר מאשר גברים, האם התגובה שלנו צריכה להיות - במחקר נבדקו סקרי הערכת הוראה פגומים, ולכן אין להתייחס לתוצאות המחקר? לדעתי לא. לדעתי אנחנו צריכים לומר - סקרי ההוראה, כפי שהם כיום, אינם מייצגים ומוטים. זהו מסר חשוב שאנחנו, המרצים, צריכים לשקף כלפי מעלה להנהלות המוסדות שלנו. האם ראוי שאני, כגבר, אקבל יתרון בתהליכי קידום משום שהסטודנטים שלי נותנים לי ציונים גבוהים יותר?"

ובמייל עוקב הוא כתב: "המחקר נראה לי ראוי מבחינה מתודולוגית (אם כי יתכן שפספסתי כשלים כלשהם). נכון, המאמר פורסם בכתב עת כלכלני, על ידי כלכלנים, אבל זה יכול להיות לכל היותר אינדיקציה לבעייתיות, לא פוסל את המאמר בפני עצמו." הקורא כתב הערה זו על בסיס פוסט 2 שבו הראיתי שחלק נכבד מהפרסומים שמוכיחים שסקרי ההוראה מוטים מופיעים בכתבי עת מתחומים ספציפיים (כמו כאן, של כלכלה). רבים מפרסומים אלו נמצאו כשגויים, קרוב לוודאי משום שהחוקרים בתחומים אלו אינם מומחים ואינם מכירים מבחינה תיאורטית ומעשית את הנושאים בבסיס סקרי ההוראה שהם הוראה, למידה, והערכת ההוראה.

הבאתי כאן את התכתובת ששימשה כבסיס לפוסט הנוכחי משום שללא ספק היא מייצגת דעות של אנשי סגל נוספים ובמיוחד של קוראים רבים של מאמרים השוללים את תקפות סקרי ההוראה, ושעלולים על סמך מאמרים מסוג זה לדרוש מהמומנים עליהם - ראשי יחידות אקדמיות - את ביטול השימוש בתוצאות של סקרי ההוראה משום שהם מוטים. להלן תגובתי למאמר.

לא קראתי את המאמר הנדון במלואו ובאופן ביקורתי כך שאינני יכולה להתייחס לאיכותו ואני מניחה לשם המשך הדיון, שהמחקר נעשה באיכות ראויה. יחד עם זאת, המסקנות מהמאמר מקיימות לפחות חמישה! מבין הפגמים שרשמתי בפוסט 1 ולפיכך לא ניתן לאמץ את ממצאיו לכלל אוכלוסיית הקורסים באוניברסיטאות ובמכללות, להלן הפירוט (ההתייחסותי המפורטת כתובה בפונט אדום):

1. **פגמים ב"בית" הפרסום וברקע של החוקרים, המציעים חוסר הבנה של התחום:** המאמר כאמור, פורסם בכתב עת בתחום הכלכלה שאינו נקרא בדרך כלל ע"י מומחים בתחום סקרי ההוראה שהם אנשי חינוך ופסיכולוגיה, כך שלא ניתנה להם הזדמנות להעביר עליו ביקורת אילו היו מעוניינים בכך וכפי שאני עושה עכשיו כשהגיע לידי. קרוב לוודאי שהוא נכתב ע"י אנשי כלכלה והוערך ונשפט לשם קבלה לפרסום גם כן ע"י כלכלנים, שאינם מומחים בתחום סקרי ההוראה ושאינן להם הבנה ראויה בנושאים מרכזיים בתחום זה.

לבעיה זו התייחסתי בפוסט 2 והיא אכן מניפה דגל אדום. יש במאמר עצמו (ראו פגמים 2-4 להלן) הוכחה ניצחת לחוסר ההבנה של כותבי המאמר בנושאי הוראה אקדמית. החוקרים מציגים בתקציר את הממצאים שלהם כאילו הם ניתנים להכללה לכלל הקורסים בכלל המוסדות האקדמיים, בשעה שהצגה זו היא שגויה לחלוטין. זאת משום שאוכלוסיית המחקר שלהם (ראו שוב נקודות 2-4) היא מאוד ייחודית ומייצגת רק אחוז קטן מאוד של הקורסים הקיימים במוסד אקדמי טיפוסי! אי לכך, תוצאות המחקר שלהם רלוונטיות רק לאחוז קטן מאוד של הקורסים ולא ניתנות להכללה לרוב הגדול של הקורסים במוסדות האקדמיים. העובדה שהחוקרים לא ציינו מידע חיוני זה בתקציר של המאמר מעידה שהם האמינו שתוצאות מחקרם ניתנות להכללה לכלל הקורסים ולכן מעידה גם על חוסר הבנתם בנושאי הוראה: הם אינם יודעים שאותו מורה מלמד באיכות שונה (וידורג בהתאם לכך בסקרי ההוראה באופן שונה) בקורסים הדורשים שיטות הוראה השונות במאוד זו מזו (כמו הרצאה לעומת תרגול) או בכיתות קטנות לעומת גדולות, ושיש שונות גדולה בהוראה בתחומי תוכן שונים.

2. **הכללה שגויה מכיתות קטנות לכלל הכיתות, כשנדרשות יכולות הוראה שונות לכיתות קטנות לעומת גדולות:** נראה שהקורא הנ"ל לא זיהה (משום שזה לא היה כתוב בתקציר של המאמר) שהמחקר נערך על אוכלוסיית קורסים קטנים בלבד שההוראה בה היא מאוד ייחודית ולכן הממצאים לא ניתנים להכללה למרבית הקורסים באוניברסיטה/מכללה: עיון קצר בתיאור השיטה במאמר מעלה שאוכלוסיית המחקר התמקדה בקורסים קטנים (עד 15 תלמידים בקורס) בעוד שבמרבית המוסדות האקדמיים בארץ (ובעולם) מרבית הקורסים הם יותר גדולים, ואפילו הרבה יותר גדולים ודורשים יכולות הוראה שונות מאשר בקורסים קטנים (למשל, פחות אינטראקציה עם הסטודנטים והוראה פעילה וממוקדת-תלמיד ויותר איכויות של הרצאה טובה והוראה ממוקדת-מורה). מחקרים מראים שאנשי סגל רבים באוניברסיטה/מכללה מלמדים קורסים קטנים וגדולים באיכות שונה (בדרך כלל אותו מרצה מדורג בסקרי הוראה גבוה הרבה יותר בכיתות קטנות מאשר בכיתות בינוניות וגדולות ולפעמים הקשר לא לינארי, ראו רשימה קצרה של מקורות למטה). לפיכך, לא ניתן להכליל מהצלחה או כישלון בהוראה בקורסים קטנים על

איכות ההוראה בקורסים גדולים יותר ולא ניתן להכליל מחקר על הוראה שנעשה באופן בלעדי על כיתות קטנות למרבית הקורסים במוסד האקדמי הנערכים בכיתות גדולות יותר.

3. **הכללה שגויה מקורסים בשיטת הוראה אחת לקורסים בשיטה שונה לחלוטין הדורשת יכולות הוראה שונות:** הקורא גם לא זיהה (שוב, קרוב לוודאי משום שזה לא היה כתוב בתקציר של המאמר) שאוכלוסיית המחקר הייתה מורים בקורסי תרגול (סקציות של שיעורי עוגן בשיטת ההרצאה) ששיטת ההוראה שבהם מאוד ייחודית והמהווים מיעוט קטן של סוגי הקורסים במוסדות האקדמיים, לעומת הרוב הגדול של הקורסים המתקיימים בשיטת ההרצאה. שיטת ההוראה השכיחה בקורסי תרגול, במיוחד בדיסציפלינות עם היבטים כמותיים כמו מתמטיקה, הנדסה, טכנולוגיה, מדעים וגם כלכלה היא שהמורים פותרים בשיעור בעיות (הרבה פעמים כאלו שניתנו כשיעורי בית), מסבירים נושאים שלא הובנו בשיעורי העוגן, משיבים לשאלות של תלמידים ומאפשרים השתתפות פעילה של התלמידים, בניגוד לאי קיומם של מרבית האלמנטים הללו בשיעורים בשיטת ההרצאה, במיוחד בכיתות בינוניות וגדולות. וכך, שיטת ההוראה בשיעורי תרגול שונה מאוד מאשר בקורסי ההרצאה ודורשת מהמורה יכולות הוראה שונות. לפיכך, תוצאות של מחקר על הוראה המבוסס רק על קורסי תרגול אינן ניתנות להכללה באופן ישיר לקורסי הרצאה.

4. **הכללה שגויה מקורסים בדיסציפלינה ספציפית יחידה לכלל הקורסים מכלל הדיסציפלינות, כשההוראה עשויה להיות שונה מאוד ולדרוש יכולות הוראה שונות מאוד:** הקורא אימץ את מסקנות המאמר היחיד הזה שאוכלוסיית המחקר שלו היא מורים בקורסים בבית הספר לעסקים וכלכלה באוניברסיטת מאסטרסכט שבהולנד לכלל הדיסציפלינות בכלל המוסדות האקדמיים ובפרט למוסד האקדמי שלו עצמו, מבלי לבדוק האם זוהי הכללה ראויה. פוסט מס' 5 שנשלח בו זמנית לפוסט זה אך בנפרד (כדי שתהיה כותרת נפרדת של הנושא) מסביר מתי ניתן להכליל ממספר קטן של דיסציפלינות לכלל הדיסציפלינות. במקרה של המחקר כאן שנערך בדיסציפלינה יחידה, לא ניתן להכליל באופן ישיר את מסקנותיו לכלל הדיסציפלינות. לדוגמה, היות והמחקר נערך על הוראה בשיעורי כלכלה בלבד, האם מובן מאליו שתוצאותיו רלוונטיים גם לקורסים באמנויות?

5. **פגמים בהתעלמות מהתיאוריות והידע העולמי הקיים בנושאי סקרי ההוראה:** הקורא אימץ את תוצאות המאמר כמות שהם, מבלי לבדוק אם נעשו מחקרים אחרים בנושא של הטיה מגדרית של סקרי ההוראה ומהם ממצאיהם פוסט מס' 6 שאשלח בפעם הבאה יציג מה ידוע ממחקרים רבים שנעשו בעולם על הטיה מגדרית של סקרי הוראה. באופן כללי נמצאו במחקרים מבוססים ואמינים גם והרבה ממצאים סותרים לממצאים במחקר זה והשאלה היא למי להאמין? על כל פנים, לפני שמאמצים את התוצאות של מחקר יחיד בנושא מסוים כאמת שאין בלתי, חשוב לבדוק מה עוד ידוע ממחקרים על נושא זה!

חמשת הפגמים המהותיים הללו מביאים למסקנה אפשרית יחידה והיא לדחות את ההכללה והאימוץ של ממצאי המאמר.

קוראים נכבדים, חשוב שתשימו לב שכל חמשת הפגמים האלו כלולים ברשימת הפגמים בפוסט 1, מה שמדגים עד כמה הרשימה שערכתי עבורכם רלוונטית ושימושית לשם גילוי בעיות במחקרים העוסקים בסקרי ההוראה (ובאופן כללי במחקרים בתחום ההוראה). אי לכך, אני פונה לכל מי שיעלה בעתיד לרשימה זו מאמרים על מחקרים המוכיחים שסקרי ההוראה מוטים, לבדוק היטב לפני כן אם אינם לקויים בפגמים הרשומים בפוסטים 1-2 ובשום אופן לא להסתפק בקריאת התקציר בלבד לשם אימוץ הממצאים!!!

רשימה קצרה של ספרות מחקר על הקשר בין גודל הכיתה ודירוג בסקרי ההוראה (לסעיף 2)

- Centra, J. A. (2009). *Differences in responses to the Student Instructional Report: Is it bias?* Retrieved from ETS, Princeton, NJ: Princeton University Press.
- Feldman, K. A. (1984). Class size and college-students evaluations of teachers and courses - a closer look. *Research in Higher Education, 21*(1), 45-116.
- Fernandez, J., Mateo, M. A., & Muniz, J. (1998). Is there a relationship between class size and student ratings of teaching quality? *Educational and Psychological Measurement, 58*(4), 596-604.
- Fourie, C. (2000). Class size and student ratings of teaching presentation. *South African Journal of Higher Education, 14*(3), 132-138.
- Gannaway, D., Green, T., & Mertova, P. (2018). So how big is big? Investigating the impact of class size on ratings in student evaluation. *Assessment & Evaluation in Higher Education, 43*(2), 175-184.
- Marsh, H. W. (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research, 11*(3), 253-388.
- Watkins, D., & Afzulpurkar, N. (1988). Class size and student-ratings of tertiary courses. *Educational and Psychological Measurement, 48*(2), 523-526