

מדדים אובייקטיביים אדישי-תוכן כנגד הערכות סובייקטיביות מושכלות

(לשם פשוטות ומתוך כניעה מבישה למוסכמות השפה העברית, השתמשתי לכל אורך הטקסט בגוף זכר רבים. כמובן שהדברים אמורים בכל חברי האקדמיה ללא הבדל מגדר. כמוכן, אינני מתייחס לשאלה לאיזו מטרה נעשית הערכת האיכות. כפי שכתבתי במקומות אחרים ישנן מטרות לגיטימיות – כגון החלטה על שכירה ומתן קביעות לאקדמאים מסוימים – ומטרות בלתי-לגיטימיות – כגון ניסיון של גורמים חוץ אקדמיים לשלוט ולנהל את האקדמיה.)

למרות שכל אנשי האקדמיה אמורים להיות מודעים למורכבות המחקר, הלימוד וההוראה האקדמיים ולמומחיות הרבה הנדרשת להבנת התכנים של עיסוקים אלו, ממשיכים רבים מהם להסתמך על מגוון של "מדדים אובייקטיביים" בבואם להעריך איכות עבודה אקדמית. מטרת המאמר הקצר הזה היא להצביע על חוסר השחר של ההסתמכות הזאת ועל הסכנה העמוקה הטמונה בה.

הדבר היחיד שניתן לומר לזכותם של "מדדים אובייקטיביים" להערכת איכות אקדמית הוא שמדדים אלו מצויים בשפע וקלים לשימוש. כמובן שזאת גם הסיבה המכרעת לשימוש הגובר בהם. אלא שהשאלה שאני מבקש להתרכז בה היא מה המדדים הכמותיים הללו מודדים. זאת שאלה מרכזית אשר המשתמשים במדדים הללו חייבים לתת עליה את דעתם **ברצינות**: זאת אומרת, לא להסתפק בתשובות סתמיות ובלתי מבוססות, אלא לבקש הוכחות לכך שהמדדים הכמותיים הללו אכן משקפים נאמנה איכות אקדמית.

לטענתי, לא רק שהמדדים הכמותיים הללו לא ממש משקפים (נאמנה!) איכות אקדמית, אלא שהשאלה הנ"ל (ז"א: השאלה "מה האיכות אשר המדדים הכמותיים הללו משקפים") לא ממש נשאלת ברצינות. המשתמשים במדדים הללו פשוט מתמכרים לאשליה שהמדדים הללו משקפים איכות, ודי בכותרת "מדדים אובייקטיביים" כדי לאפשר את האשליה הזאת.

האשליה של מדדים אובייקטיביים

מה שעומד בבסיס האמון במדדים האובייקטיביים הוא אופיים המתמטי והבלתי-אישי. בפרט, שימוש חוזר באותו מדד לגבי אותה שאלה ייתן את אותה תשובה בלי קשר לזהות המשתמש ומיקומו. זאת כמובן בתנאי שהמשתמש יודע איך להשתמש נכון במדד, תנאי סובייקטיבי שמשום מה לא נותנים עליו מספיק את הדעת.

אבל גם הגובה של חברי האקדמיה הוא מדד אובייקטיבי, כך גם משקלם, גילם וכו'. באופן דומה, גם האורך הכולל של פרסומיהם הוא מדד אובייקטיבי, וכך גם משקלם של הספרים על מדפי ספרייתם הפרטית, ממוצע גילי הסטודנטים שלהם, וכו'. ובאופן דומה, גם מספר המאמרים שהם פרסמו ומספר הציטוטים הם מדדים אובייקטיביים. כמובן שיש הבדל: עד כמה שאני יודע אף אחד לא משתמש במדדים הראשונים בעוד שרבים משתמשים באחרונים, ובעוד שהקשר לאיכות אקדמית נראה מופרך כאשר מדובר במדדים הראשונים, הוא נראה סביר כאשר מדובר באחרונים. **אבל האם מראית העין הזאת נכונה? יתר על כן, האם הקשר הזה הוכח באופן משכנע כל כך שניתן להסתמך עליו בקביעת גורלות של אקדמאים ספציפיים ושל יחידות אקדמיות שלמות** (שלא לציין את הדיבורים על האיכות של אוניברסיטה שלמה או של אוסף האוניברסיטאות בארץ זו או אחרת)?

אני טוען שהקשר בין המדדים הכמותיים השונים לבין איכות אקדמית לא נבדק **ברצינות** מעולם. כמובן שהאנשים שפיתחו את המדדים האלו ניסו, כמיטב הבנתם, ליצור מתאם בין המדדים לבין האיכות. אבל הבנתם של מפתחי המדדים הנ"ל רחוקה מלהיות מספקת, פשוט משום שהם אינם מומחים (ואינם יכולים להיות מומחים) בכל תחומי הידע האקדמיים. אבהיר את הטעון בפסקה הבאה, אבל לפני כן אני מבקש להצביע על העובדה שגם אם מפתחי המדדים האלו היו מוכשרים למשימה המדוברת, הרי המדד שהם היו יוצרים היה משקף את ההבנה הסובייקטיבית שלהם בתחומי הידע האקדמיים השונים, ולפיכך לא היה ראוי להחביא את הסובייקטיביות הזאת תחת הכותרת "מדד אובייקטיבי".

הטענה שמדד מסוים משקף איכות מסוימת היא טענה המתייחסת לתוכן מסוים, ובחינתה היא עניין למומחים בתוכן המדובר. פיתוח ובחינה של מדד המשקף מחלות מסוימות הוא עניין למומחים ברפואה, ופיתוח ובחינה של מדד המשקף מיומנות בתחום טכנולוגי מסוים הוא עניין למומחים לטכנולוגיה הזאת. בשני המקרים אי אפשר להסתמך רק על "מומחים לבני אדם" או על מומחים לפיתוח מדדים. באופן דומה, מדד של איכות אקדמית בתחום דעת מסוים צריך להבחן על ידי מומחים בתחום זה, ולא על ידי "מומחים לאקדמיה" שאינם מבינים את התכנים המקצועיים של אותו תחום וגם עשויים שלא להכיר ולהבין את התרבות האקדמית שלו. החשיבות של התכנים המקצועיים ושל התרבות האקדמית תודגם בהמשך.

כפי שנרמז לעיל, פיתוח של מדד כמותי ("אובייקטיבי") של איכות מורכב משני חלקים. החלק הראשון והמכריע, הוא קביעת קריטריונים: זאת אומרת, מה מודדים, ואיך ממשקלים את המדידות השונות. הקביעות הללו מבוססות על שיקול דעת סובייקטיבי (ולכן טענת ה"אובייקטיביות" של המדד היא תמיד אשליה), וכטיב שיקול הדעת – טיב המדד. לדוגמא, אם רוצים לבסס את הערכת האיכות האקדמית על מדד של פרסומים, צריך לדעת עד כמה פרסום בכתב עת מסוים משקף איכות אקדמית. למומחים בתחום הידע המסוים יש יתרון מכריע בהערכת השאלה הזאת, אבל בהמשך אטען שגם הם אינם יכולים לפתח מדד עם תואם איכות מספק (משום שמגוון איכות המאמרים בכל כתב עת הוא עצום).

החלק השני של המדד הוא איסוף אינפורמציה ועיבוד סטטיסטי שלה. אבל, גם כאן, בהקשר של מדדים המתיימרים לשקף איכות אקדמית, מדובר באשליה של אובייקטיביות: אם מדובר בספירת פרסומים הרי שהם מסתמכים על חוות הדעת של מספר קטן (בד"כ שנים או שלושה) של שופטים אנונימיים, שאיש פרט לעורכים אינו יודע את טיב שיקול הדעת (הסובייקטיבי) שלהם. זאת אומרת, לא רק שמדובר בגורם סובייקטיבי, אלא שלמשתמש במדד אין מידע על מידת מומחיותו ומהימנותו של גורם זה. ואם מדובר על ספירת ציטוטים הרי שגם כאן מדובר בהחלטות סובייקטיביות שהמניעים להן אינם ידועים. (אפשר כמובן לטעון שכל המניעים הסובייקטיביים הללו "מתקזזים" כאשר מסכמים מספר רב של פריטים, אבל האם מספר הפרטים המסוכמים גדול מספיק? ומה מידת המתאם המושגת?)

אולם החשש המרכזי שלי מתייחס לקביעת הקריטריונים. המדדים של פרסומים וציטוטים מתייחסים לכתבי-עת מסוימים ולעתים כוללים שקלול שלהם. מי קובע את רשימת כתבי-העת הנלקחים בחשבון ומי משקלל אותם ועל סמך אילו שיקולים? לפחות בתחום המומחיות שלי (מדעי המחשב התיאורטיים), השיקולים העומדים בבסיס כמה מן המדדים המשמשים כיום להערכת איכות אקדמית הינם שגויים באופן קיצוני. מצב זה גורם לכך שאין, לדוגמה, שום מתאם בין ה"אימפקט פקטור" של כתבי עת שונים לבין המעמד שלהם בקהילה המדעית. כמובן, התרבות האקדמית של תחום המומחיות שלי מחשיבה פרסומים בכנסים (בעיקר בכנסים המרכזיים) יותר מאשר פרסומים בכתבי-עת, אבל נראה שמפתחי המדדים לא היו מודעים לכך (או שלא נתנו לכך משקל מספיק). בתחומים אחרים, בעיקר בתחומים אשר עניינם חקר החברה המקומית, יש כתבי-עת מקומיים בעלי חשיבות רבה, ואלו בדרך כלל אינם "נספרים" במדדים השונים.

הדוגמה האחרונה מראה שקביעת קריטריונים (של "מדד איכות") יכולה להיות קשורה באופן עמוק בהשקפת עולם אקדמית: לדוגמה כמה חשיבות מייחסים למחקר של תופעות מקומיות ועד כמה חזקה ההעדפה של תופעות "אוניברסאליות" (אשר הכופרים במושג האוניברסאליות רואים כתופעות "מערביות" ו/או "אנגלוסקסיות"). ההכרעה בשאלה של קריטריונים למדדי פרסומים כרוכה כאן בשאלה העקרונית של תכני המחקר, ומכאן הסכנה שבהפקרתה למפתחי מדדים כמותיים.

אם נפנה לדוגמה אחרת, הרי שמדדים בינלאומיים שונים של דירוג אוניברסיטאות נותנים משקל שונה לכמות הסטודנטים "הזרים" באוניברסיטה. מה המשקל הנכון? מה הקביעה הזאת משקפת? (מיותר לציין שמשקל גבוה יוריד את דירוג האוניברסיטאות הישראליות ויעלה מאוד את דירוג האוניברסיטאות האמריקאיות – אגב, היכן יושבים מפתחי המדדים הללו?)

לקראת סיום פרק זה, אחזור לטענתי שגם מומחים בתחום ידע מסוים לא יצליחו ליצור מדד כמותי עם מתאם גבוה (במידה מספקת) לאיכות אקדמית. הסיבה היא שאפילו המחקר בתחום ידע מצומצם ביותר הוא רב-גוני מכדי שאפשר יהיה לעשות לו רדוקציה למדד מספרי מן הסוג המדובר. בפרט, בניגוד לאמונות נאיביות, אפילו באותו כתב-עת מתפרסמות עבודות אשר שונות זו מזו באופן משמעותי באופיין, נושאן, היקפן, עומקן, חשיבותן המיידית, השפעתן לטווח ארוך, ובסיכום אחרון – באיכותן (אשר בלתי אפשרי לפתח נוסחא לחישובה כפונקציה של הגורמים הנ"ל).

אדרבה, יעלו הקוראים בדעתם מבחר של כמה עשרות עבודות מחקר בתחום מומחיותם וינסו ליצור מדד אדיש-תוכן בעל מתאם מספק לאיכות עבודות אלו. מתאם טוב מספיק צריך לא רק להפריד בין עבודות המופת לבין עבודות רגילות, אלא צריך לתת דירוג טוב בתוך כל קטגוריה ובפרט להבחין בין עבודות משפיעות יותר למשפיעות פחות. המתאם חייב להיות טוב מספיק על מנת שאפשר יהיה לבסס עליו החלטות של קביעות ו/או טענה שמחלקה אחת טובה יותר ממחלקה אחרת, וכן הלאה. אני, למרות מומחיותי בתחום המחקר שלי והיותי בעל כישורים מתמטיים וניסיון רב בבניית מודלים, אינני יכול למצוא מדד כזה לתחום המומחיות שלי.

לסיכום פרק זה, אציין שלא רק ש"המדדים הכמותיים לאיכות אקדמית" אינם מהימנים, אלא שאשליית האובייקטיביות שלהם הינה מסוכנת ביותר. ראשית, משום ש"האובייקטיביות" הזאת מחביאה עמדות עקרוניות הנתונות במחלוקת (וטוב שהן נתונות במחלוקת, אך חשוב שהמחלוקת תהיה גלויה), ושנית משום שהן מעודדות הזנחה של הערוץ המהימן היחיד להערכת איכות אקדמית: חוות דעת המבוססות על הבנה ושיקול דעת.

חוות דעת סובייקטיביות המבוססות על הבנה ושיקול דעת

הצורה היחידה להעריך, באופן מהימן, איכות אקדמית היא להתבסס על חוות דעת של מומחים בתחום המדובר. רק חוות דעת כאלו יכולות להתייחס לתוכן של עבודות מחקר, ורק על בסיס התוכן של עבודות מחקר ניתן להעריך את איכותן.

אכן, מדובר בחוות דעת סובייקטיביות בעליל, אבל הסובייקטיביות משקפת הבנה ושיקול דעת שהינם נחלת חוקרים מסוימים בתחום הדעת המסוים, ואינם נחלת חוקרים אחרים (או שהוא נחלתם במידה פחותה). כמובן שגם את המומחים הללו אי אפשר לאתר באמצעות מדדים אובייקטיביים, גם לצורך האיתור הזה נחוצה הבנה ונחוץ שיקול דעת.

מה לעשות, החיים – ובעיקר החיים האקדמיים – מסובכים: אין תחליף להבנה ושיקול דעת. אלא שאין סיבה להיבהל: המשימה של איתור מומחים המצטיינים בהבנה ושיקול דעת אינה בלתי אפשרית – כך התנהלה האקדמיה במהלך מאה השנים האחרונות.

כמובן שההתבססות על חוות דעת סובייקטיביות אינה חופשייה מקשיים וסכנות. אבל קל להתמודד עם הקשיים והסכנות הללו כאשר התוכן גלוי ולא מוסתר תחת "מדדים אובייקטיביים" (שכאמור, אופן קביעתם סמוי מן העין). אפשר וצריך להפעיל שיקול דעת בבחירת נותני חוות הדעת, ואפשר וצריך להפעיל שיקול דעת בקריאת חוות הדעת שלהם. כל זה מחייב התייחסות לתוכן והבנה של התוכן. וגם כאן, הדברים סבוכים מכדי שאפשר יהיה לתת מרשם לביצוע (ובפרט לביצוע אוטומטי ו/או אדיש-תוכן).

הנ"ל מתייחס למקרה המצוי שבו יש כוונה כנה לברר את האיכות האקדמית של עבודות מחקר ו/או חוקרים. כמובן שבכל מערכת אנושית יש ניסיונות להטיה ואפילו לרמאות. את הניסיונות האלה יש לבלום תוך שימוש באותם כלים: השגת חוות דעת ממומחים המצטיינים בהבנת התחום, בשיקול דעת, וביושר. כמובן, מי שמבקש לפקח על תהליכי הערכת איכות בתחום אקדמי מסוים, חייב לרכוש הבנה בתכנים של תחום זה ולהכיר את התרבות האקדמית שלו.

כאשר מדובר בהערכת איכות של פעילות אקדמית, חשוב להבין שיש בעיה מהותית בסתירה שבין העובדה שהבנה עמוקה של התכנים האקדמיים מחייבת קירבה לתחום, ובין החשש שהקרובים לתחום עשויים להעריך את תכניו ו/או חוקריו באופן חיובי יותר מן הראוי. כמו בהרבה מקרים אחרים, הפתרון לבעיה מצוי בדרך האמצעי: מערכת ההערכה צריכה לכלול (ואכן כוללת) מגוון של אקדמאים, חלקם קרובים יותר לתחום וחלקם קרובים פחות, אבל עדיין כאלה המסוגלים להבין את תכניו של התחום ואת אופי הפעילות בו. הקרובים יותר מספקים חוות דעת מקצועיות, והקרובים פחות קוראים, מקשיבים, ושואלים שאלות בתהליך הדומה לחקירת שתי וערב בבית משפט. אגב, גם כאן, המטרה היא בירור האמת, וגם כאן עדות מוחזקת כנכונה עד שהוכח אחרת.

התהליך הזה חוזר על עצמו בכמה רמות: גם ברמת המחלקה ישנם קרובים יותר וקרובים פחות, וודאי שכך ברמות גבוהות יותר (לדוגמה: ועדת המינויים העליונה של המוסד). אבל חיוני שהמשתתפים בכל אחד מן השלבים יזכרו מי המומחים ומי המפקחים: זאת אומרת, בידי מי מצויה ההבנה העמוקה יותר של התכנים. גם חשוב שהמפקחים יבינו את התרבות האקדמית של התחום וייתחסו לאמות המידה שלו; זאת אומרת, לא ינסו לכפות על תחום דעת אחר את אמות המידה הטכניות של תחום המומחיות שלהם.

עודד גולדרייך, פרופסור למדעי המחשב, מחזיק הקתדרה ע"ש מאיר וייסגל, מכון ויצמן למדע