

מדדי איכות אדישי-תוכן ומשטור האקדמיה¹

עודד גולדרייך, הפקולטה למתמטיקה ומדעי המחשב, מכון ויצמן למדע

תקציר

המאמר עוסק בפרקטיקה של הסתמכות על מדדים כמותיים אדישי-תוכן לצורך הערכת איכות של מחקר אקדמי. הוא נפתח בהגדרה של המדדים הללו תוך הבחנה בין המידע הביבליוגרפי הבסיסי לבין אופן עיבודו. לאחר מכן נחשפות ההחלטות הסמויות אשר קובעות את המידע הביבליוגרפי הבסיסי ואת אופן העיבוד הסטטיסטי שלו. בפרט, מדד כמותי אדישי-תוכן מתקבל על ידי עיבוד אוטומטי של פרמטרים שטחיים של פרסומים הנובעים ממחקר, והיומרה האובייקטיבית של המדד מסתירה את השרירותיות של ההחלטות שעליהן הוא מתבסס ואינה מאפשרת דיון בהן.

ההשלכה הישירה של השימוש הדומיננטי במדדים אדישי-תוכן הינה החלטות בלתי מושכלות (הלכה ולמעשה), משום ששיקול דעת מקצועי אשר מתייחס לתוכן מוחלף על ידי מדד שטחי שהרלוונטיות שלו להחלטות שעל הפרק מוטלת בספק. ההשלכות העקיפות של השימוש במדדים אלו הינן מסוכנות עוד יותר: מדובר בהזנחת התוכן ובפטישיזציה של הכמות, בהחלפת תודעת היעוד האקדמי על ידי התכוונות לביקורות חשבונאיות, בדיכוי הסקרנות האינטלקטואליות, בעידוד מניפולציות בלתי-ענייניות, ובשכפול יחסי הכוחות והגדלת האי-שוויון בתוך תחומי ידע וביניהם. המדדים הכמותיים משתלבים היטב בשינוי מטרות ההערכה מבדיקה של עמידה בתנאי סף לצרכים פנימיים ליצירת של דירוג לצרכים חיצוניים.

המאמר מסתיים בכמה הסברים לפופולריות של המדדים הנ"ל: השתלבות בסדר הניאו-ליברלי על מנגנוני הרגולציה האחידים שלו, הפיתוי האובייקטיבי (והתאמתו לרוח המדע המודרנית), אופורטוניזם (עצלנות מחשבתית והתנערות מאחריות), אמצעי שליטה של המעמד הבירוקרטי ("המנהלים") במעמד החוקרים בפועל ("העובדים") אשר מאפשר הכפפה של ידע מדעי לידע ניהולי, דיפוזיה של הגישה העסקית אל השדה האקדמי, והשתלטות של טכנולוגיה (ביבליומטרית) על תוכן מדעי. כל אלה קשורים לעלייה של משטר פוליטי-מדעי חדש אשר משטור הדוק יותר של האקדמיה הוא חלק מהותי ממנו.

מאמר זה הינו מאמר דעה. הוא מתאר ומנתח תופעה מציאותית, תוך נקיטת עמדה ברורה לגביה. העובדות הבסיסיות אינן נתונות במחלוקת, אולם ההערכה של מידת השפעתן והיחס כלפי השפעה זו נתונים במחלוקת. התופעה עצמה היא גלובלית ומקומית, והניתוח יתמקד באוניברסלי תוך הדגשת המקומי לפי הצורך.

¹ לשם פשטות ומתוך כניעה מבישה למוסכמות השפה העברית, השתמשתי לכל אורך הטקסט בגוף זכר רבים. כמובן שהדברים אמורים בכל חברי האקדמיה ללא הבדל מגדר.

1. הקדמה

למרות שכל חברי האקדמיה אמורים להיות מודעים למורכבות העיסוקים האקדמיים (קרי: מחקר והוראה) ולמומחיות הרבה הנדרשת להבנת התכנים של עיסוקים אלו, קל וחומר להערכת איכותם, נוהגים רבים מהם להסתמך על מגוון של מדדים כמותיים אדישי-תוכן של איכות אקדמית. מה שעומד בבסיס האמון במדדים הללו הוא אופיים המתמטי והאובייקטיבי (או הבלתי-אישי). בפרט, שימוש חוזר באותו מדד לגבי אותה שאלה ייתן את אותה תשובה, בלי קשר לזהות המשתמש ומיקומו.²

אך 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. כמוכן שיש הבדל: 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)?

השאלות הנ"ל אכן מתבקשות. אלא שגם הן לוקות בחסר משום שהן מניחות שידוע ומוסכם מהי איכות אקדמית, שידוע ומוסכם שיש להעריך אותה (בפרט באותם מקרים בהם מבקשים לעשות זאת), ושתהליך ההערכה מבקש לתאר מצב (ולא ליצור מצב). ההנחות הללו הינן נאיביות, וראוי לתהות בביקורתיות על מה שהן מניחות כמוכן מאליו. בפרט, ראוי לפתוח בשאלות הבאות.

- מהי איכות אקדמית? מי קובע מהי איכות אקדמית ולשם מה?
- לשם מה ומתי יש לבצע הערכה של איכות אקדמית?
- האם הערכת איכות מתכוונת רק לאמוד מצב נתון וקיים או ליצור מצב חדש? במקרה שמדובר בניסיון ליצור מצב חדש (או "רק" להשפיע על המצב הקיים), מהו המצב החדש שמנסים ליצור ומה ההצדקה לכך? ואיך מתייחס השינוי הזה לאיכות האקדמית שעליה נסובה השאלה הראשונה?

מאמר זה לא מתיימר לתת הגדרה לאיכות אקדמית, אך עקרון יסודי לכל הגדרה סבירה של איכות הוא **שאיכות של פעילות בשדה מסוים היא דבר פנימי לשדה אשר ניתן להערכה רק מתוך השדה**.³ הערכה של פעילות בשדה דורשת הבנה טובה של השדה – כלומר, מומחיות בשדה עצמו. מומחיות מלאה או חלקית ניתנת לרכישה, גם על ידי גורמים חיצוניים לשדה, אבל חייבים לזכור כי רמה המומחיות בשדה קובעת את רמת המהימנות של הערכת האיכות של פעולות בשדה. בפרט, לצורך הערכה של פעילות בשדה מסוים, לא ניתן להחליף הבנה של השדה המסוים בהבנה כללית של הדינמיקה של שדות.⁴

² זאת כמוכן בתנאי שהמשתמש יודע כיצד להשתמש נכון במדד, תנאי בלתי-אוניברסלי בעליל שמשום מה לא נותנים עליו מספיק את הדעת. כמוכן שהערה זאת נאמרת בבדיחות הדעת. כך גם לגבי תחילת הפסקה הבאה.

³ ניסוח אחד, מבין רבים, לעקרון זה מצוי אצל מקינטאיר (2006) בפרק "טבע המידה הטובה".

⁴ המונח "דינמיקה של שדות" מושאל מבורדייה, אך בהקשר הנוכחי הוא מוחל גם על ידע כללי ביחס למצב הנוכחי של האקדמיה והדיסציפלינות השונות. זהו סוג הידע שעליו מבוססים המדדים אדישי-תוכן הנפוצים ביותר, והוא עומד בניגוד להבנה מעמיקה ורחבה של התוכן הנוכחי של דיסציפלינה מסוימת.

כל העקרונות הנ"ל, אשר נגזרים מעקרון היסוד הראשון (ז"א הבנה של "איכות" כמהות תוכנית-פנימית לשדה), פוסלים את הניסיון להעריך איכות אקדמית של פעילות בדיסציפלינה מסוימת על סמך מדדים כמותיים אדישי-תוכן וללא כל הבנה של התוכן של הדיסציפלינה הזו. ניתוח מפורט יותר של חוסר המהימנות של הניסיונות הללו עומד במוקד המאמר.

עניין מרכזי נוסף שיש לתת עליו את הדעת הוא העיסוק האובססיבי בהערכת איכות אקדמית במקום בניסיונות כנים לתרום לשיפורה. ההנחה המובלעת בעיסוק הזה היא שהערכת איכות מובילה לשיפור באיכות, וזאת על ידי מגננים של תגמול וענישה המוזנים על ידי הערכת האיכות. הנחה זאת ניתנת להצדקה רק במצבים מסוימים, ובמצבים אחרים – שבהם מתקיים עיסוק יתר בהערכת איכות – התוצאה הינה פגיעה באיכות ולא שיפור שלה.⁵ בכל מקרה, עיסוק בהערכת איכות לא יכול להחליף הקצאה של משאבים. למותר לציין, שיש מצבים בהם **איכות נמוכה דורשת תוספת של משאבים ולא גריעה שלהם.**

עיסוק היתר בהערכת איכות אקדמית קשור באופן הדוק בהופעת "תרבות הניטור" (Audit Culture) כחלק מן ההגמוניה הניאו-ליברלית (Shore 2008). הכוונה היא להשתלטות של מנטליות חשבונאית על אופן הארגון החברתי ובפרט על אופן הניהול של מוסדות ממשלתיים וציבוריים. הניטור החשבונאי, אשר מופעל בדרך כלל מחוץ לשדה המוערך, גורם להתעלמות מן התוכן של הפעילות בשדה ולהתכחשות למשמעות שלה. בהקשר של האקדמיה, תופעות אלו מגיעות לשיאן עם עליית השימוש במדדים כמותיים אדישי-תוכן, והתוצאה הכמעט מיידיית הינה **פיחות בחשיבות המיוחסת לפעילות עצמה והתרכזות בשיפור ערך המדד.** בטווח ארוך יותר, התוצאה הינה זניחת המטרות המהותיות של המערכת האקדמית – אשר הינן טרנספורמטיביות ביחס לחברה – והחלפתן במטרות שמרניות של קיום ושכפול הסדר החברתי הקיים.⁶ אכן, יש אירוניה בכך שהרצון לשמר את הסדר החברתי הקיים מוביל לרצון לשנות את האקדמיה.

יש קשר מהותי בין הניסיון לבצע הערכת איכות חיצונית של פעילות בשדה לבין הדומיננטיות של מדדים כמותיים אדישי-תוכן בהערכה זאת. הרדוקציה של תוכן לכמות מבטלת את היתרון המובנה של מומחיות בשדה, שהרי הכמות נהירה באופן זהה לכולם. יתר על כן, נוצר יתרון מובנה לאלו אשר מרגישים בנוח עם התעלמות מתוכן השדה ועם התייחסות לפעילות בשדה דרך הפריזמה של מדדים כמותיים אדישי-תוכן. כך מתאפשר דיכויים של העוסקים בפעילות בתחום האקדמי (ז"א: העובדים האקדמיים) ומשטורם ע"י שדרת הניהול האקדמית (והמעמד אשר שולט בפוליטיקה החיצונית).

שתי הפסקות האחרונות מבהירות שבניגוד להעמדת הפנים שעל פיה תהליך הערכת האיכות משקף ו/או מתאר מצב קיים, הרי שלמעשה תהליך הערכת האיכות (החיצוני לשדה) הוא תהליך אשר מכוון מצב חדש: העובדה שתהליך הערכת האיכות משנה את הפעולות והתכנים בתחום המוערך אינה תקלה שמקורה במימוש מסוים של התהליך, אלא היא חלק מהותי של התהליך. כאמור לעיל, מדובר בשינוי מרחיק לכת של מטרת העל של המערכת האקדמית, ממערכת שה**אתוס** המרכזי שלה הוא שינוי החברה באמצעות גילוי והפצה של ידע ודרכי המשגה חדשים ומעשירים, למערכת שמכוונת במוצהר לשימור

⁵ ניתוח ההשפעות של עיסוק יתר בהערכת איכות על החינוך הטרנס-אקדמי מופיע בספר על מערכת החינוך בפילנד (סאלברג 2015). בפרט, מדובר על האפקט ההרסני של התרכזות בשיפור הדירוג ביחס למבחנים חיצוניים וסטנדרטיים, אשר זרים למטרות המערכת ואדישים לשונות של אוכלוסיית הלומדים, כאשר התרכזות זו באה על חשבון לימוד והתעמקות בתוכן עצמו. אספקט הרסני נוסף הוא פגיעה קשה ביצירתיות ובמוטיבציה של המלמדים (קרי: המורים) שהם המשאב העיקרי של מערכת חינוך טובה. דברים אלו תקפים קל וחומר כאשר מדובר במערכת האקדמית.

⁶ דוגמאות קונקרטיות לשינוי מטרות ופעולות בתגובה לניטור יתר (בפרט, תוך הסתמכות על מדדים כמותיים אדישי-תוכן) מופיעות במאמר של אשפלנד ואסאודר (Espeland and Sauder 2007) אשר מתמקד במקרה המבחן של קבלה לבתי ספר למשפטים בארה"ב. לדיון רחב יריעה בשינויים במערכת האקדמית, ראו Halffman and Radder (2015).

ולשכפול הסדר החברתי הקיים, מגמה המגיעה לשיאה בתפיסת הסטודנטים כרוכשים "מיומנות מקצועית" גרידא. העובדה שגם בעבר המערכת לא הייתה מחויבת במידה מספקת לאתוס שלה אינה סותרת את האבחנה כי התהליך הנוכחי מוביל לזניחה מוחלטת של האתוס הזה.

1.1 בין איכות, להשפעה ונראות

בעוד שהפער המושגי בין איכות לבין כמות בולט לעין, קל יותר להתבלבל בין **מהות תכנית** לבין מדד כמותי אדיש-תוכן כאשר מתייחסים למושגים כגון השפעה (impact) של מחקר או חוקר, ונראות (visibility) שלו. אולם גם ביחס למושגים אלו, שבמהותם התכנית הם מומנטים של איכות, מתקיים פער מושגי בין המהות התכנית לבין המדד הכמותי.

כאשר דנים בהשפעה של מחקר מסוים, הכוונה היא להשפעה של אותו מחקר על תחום הידע הרלוונטי כפי שהיא באה לידי ביטוי בתרומה שלו לתמונת העולם של המומחים בתחום, תרומה אשר במובן מסוים מחוללת שינוי – למצער קטן – של תמונת העולם שלהם. התרומה עשויה להיות תוספת של מידע עובדתי ו/או פרשנות מושגית חדשה של המידע הקיים או ערעור על מידע עובדתי המוחזק כנכון ו/או על פרשנות מושגית ידועה. תרומה כזו יכולה להיות משמעותית יותר או פחות, בעלת תחולה רחבה יותר או פחות, בעלת פוטנציאל מפרה יותר או פחות, וכו'. מכאן שההשפעה במובן המהותי (הזה) כרוכה בתוכן של תחום הידע הרלוונטי, והערכתה מחייבת הבנה של התוכן הזה. המובן המהותי הזה הוא המובן הקשור לאיכות המחקר ולתרומתו לתחום, ויש פער מושגי וממשי בינו לבין מספר הציטוטים של פרסום מסוים. בפרט, מספר הציטוטים אינו מעניין כשלעצמו אלא כאינדיקציה להשפעה המהותית של המחקר, אלא שזו אינדיקציה מאוד לא מהימנה (מסיבות שיפורטו בהמשך המאמר): קורלציה חזקה של המדד הכמותי עם השפעה במובן המהותי מתקיימת רק במקרים קיצוניים, בעוד שהשימוש הנפוץ במדדים הוא במקרים רגילים.

לרוע המזל, השפעה במובן המהותי ניתנת לבלבול בקלות עם השפעה במובן הכמותי הבאה לידי ביטוי במדדים ששמן מכיל את המילה "אימפקט" (impact) אשר מתעלמים מן התוכן עצמו. באופן דומה, "נראות" מתייחסת במהותה להערכה שחוקר מסוים זוכה בתחום מסוים, אך קל לבלבל בינה לבין מדדים כמותיים של הופעות של החוקר בפורומים מסוימים. בשני המקרים, ההשפעה והנראות המהותיים מתייחסים להערכה של החוקרים בתחום (ז"א רפלקציה פנימית לתחום), בעוד שהמדדים הכמותיים אשר ממותגים בשמות אלו מייצגים הסתכלות חיצונית לתחום (אשר מתייחסת למראית העין ולא לתוכן עצמו). מכיוון שהמשמעות המהותית של המושגים של השפעה ונראות קשורה למושג האיכות האקדמית, בחרתי לנהל את הדיון תוך התייחסות למושג האיכות (ולא במונחים של השפעה או נראות), וזאת משום שלגבי מושג זה קיים פחות בלבול מושגי (בין המהות הנטועה בתוכן מקצועי לבין מה שנמדד ללא התייחסות לתוכן עצמו).

1.2 פרקי המאמר

להלן תיאור קצר של פרקי המאמר. ראשית (בפרק 2) נאפיין את המושג של מדד כמותי אדיש-תוכן, נתאר מספר סוגים שכיחים של מדדים כאלו, ונסקור בקצרה את ההיסטוריה של פיתוחם והשימוש בהם. משם נעבור (בפרק 3) לחשיפת שורה ארוכה של החלטות בעייתיות אשר מסתרות תחת החזות של מדד אובייקטיבי. החלטות אלו כוללות את עצם ההחלטה להתעלם מן התוכן של הפרסום ולהגדיר מאמרים המתפרסמים באותה זירה כשוויים, את ההחלטה המגדירה מה נחשב פרסום מדעי ובהתאם לכך מה יכלל במסד הנתונים הביבליוגרפיים, ואת בחירת העיבוד הסטטיסטי המסוים המופעל על מסד הנתונים הנ"ל (ז"א קביעת הנוסחה המתמטית אשר מייחסת ערך מספרי מסוים לאוסף מאמרים מסוים). נשאל מי מחליט לגבי שאלות אלו, על סמך אילו שיקולים, ונצביע על היחס בין החלטות אלו לבין יחסי הכוחות שבין תחומי ידע שונים ובתוכם.

כפי שנרמז לעיל, המדדים הכמותיים אינם "ניטרליים" אלא מבטאים היררכיות בין תחומי ידע ובתוכם, היררכיות שהינן מקובלות על מפתחי המדדים. בפרט, מדדים אשר מפותחים על ידי מדענים ומהנדסים

אשר קרובים לתחומי ידע מסוימים נוטים להתאים לתרבות הפרסומים בתחומים אלו, ועלולים להיות מאוד בלתי-מתאימים לתחומי ידע בהם הנורמות הפרסומיות שונות. מכאן **השונות הלגיטימית שבין תרבויות מקצועיות שונות אינה מכובדת**, משום שהמדדים מתאימים לנורמות פרסומיות אשר נפוצות בתחומי ידע מסוימים אך לא באחרות.⁷ הדבר בא לידי ביטוי באופן הבטה ביותר בהטיה הברורה של המדדים הנפוצים לטובת זירות פרסום בין-לאומיות (ובפרט אנגלו-סקסיות) ולעיתים בהתעלמות מוחלטת מזירות פרסום מקומיות, ואזריות ו/או לאומיות. מיותר לציין שאם המדדים משפיעים על הקצאת משאבים, הרי שהם **יוצרים העדפה ברורה של תחומי ידע מסוימים ודיכוי של תחומי ידע אחרים**. בפרט, אחת התוצאות היא הזנחה של מחקרים הממוקדים בחברה ובתרבות המקומית – דבר **שהרסני במיוחד בתרבות פריפריאליות כמו החברה הישראלית**.

בנקודה זאת נעבור להשלכות הישירות והעקיפות של שימוש דומיננטי במדדים כמותיים אדישי-תוכן לשם הערכת איכות מדעית (פרקים 4 ו-5 בהתאמה). **ההשפעה הישירה של שימוש כזה היא קבלת החלטות בלתי-מושכלות** הלכה למעשה: להלכה משום שמעצם הגדרתם המדדים מתעלמים מן התוכן אשר אמור להיות הבסיס להחלטות מושכלות, ולמעשה משום שהמתאם בין המדדים לבין האיכות הינו בדרך כלל נמוך מדי מכדי להצדיק הסתמכות על המדדים. בפרט, נצביע על השונות הרבה שבין איכות מאמרים, ובפרט גם בין איכות מאמרים המתפרסמים באותה זירה, ועל השונות הרבה שבין חשיבות המאמרים המצוטטים באותו מאמר (מבחינת תרומתם של המאמרים המצוטטים למאמר המצוטט, קל וחומר מבחינת תרומתם לתחום הידע בכללותו). נדגיש שתופעה זאת מתקיימת גם בתוך תחום ידע יחיד שכל העוסקים בו שותפים לאותה תרבות מקצועית; כמובן, שהיא חזקה עוד יותר כאשר מדובר בפרסומים המתאיחים לכמה תחומי ידע שונים (דבר שמוסיף בעיתיות במקרים של פרסומים אינטר-דיסציפלינריים).

ההשפעה הישירה של שימוש דומיננטי במדדים כמותיים אדישי-תוכן מוחמרת ע"י נטייתם **להסתיר את ההחלטות** הבעייתיות העומדות בבסיסן ולמנוע דיון בעמדות העקרוניות שהובילו להחלטות אלו. ההעדפה של פרסומים מסוג מסוים מעוגנת תמיד בעמדות לגבי אופי המחקר הראוי בתחום הידע הרלוונטי, ועמדות אלו נתונות במחלוקת כמעט בכל תחום. כמובן שמדובר במחלוקות לגיטימיות ומפרות שראוי שתהיינה פתוחות לדיון, אלא שהמדדים הכמותיים מבוססים על עמדה אחת, אינם מכריזים עליה במפורש, ובכך מונעים דיון בה. בנוסף לכך, ההסתמכות הגוברת על מדדים כמותיים מובילה להזנחת הערוץ המהימן היחיד להערכת איכות אקדמית: חוות דעת המבוססות על הבנה של תוכן הפעילות המוערכת ושיקול דעת המבוסס על מומחיות בתחום הידע הרלוונטי.

הנזקים העקיפים של השימוש הדומיננטי במדדים כמותיים אדישי-תוכן חמורים עוד יותר מן הנזקים הישירים. הנזק הברור ביותר הוא בדחיפה נוספת של האקדמיה ממחויבות למחקר כנה ואמיץ, שסימנו המובהק הוא האיכות של התוכן, להתעסקות בכמות של פרסומים. כמובן שמטרת המחקר היא לא רק לגלות ידע חדש או צורת הסתכלות חדשה אלא גם לשתף את הקהילה המדעית הרלוונטית בידע הזה, והשיתוף הזה נעשה דרך המדיום של פרסומים. עובדה זאת לא משנה את הראיה של המחקר כמטרה ואת הפרסום כאמצעי לשיתוף. הדומיננטיות של המדדים הכמותיים אשר אדישים לתוכן מערערת את הראיה הזאת: באווירה הנוצרת, **הפרסומים הופכים למטרה, ואילו המחקר נחווה כאמצעי** להשגתם.

הפיכת הפרסומים האקדמיים מאמצעי למטרה אינה חידוש של המדדים הכמותיים, אלא תהליך היסטורי שמלווה את הפיכת המחקר האקדמי מתחביב אריסטוקרטי למקצוע (יוקרתי). לפיכך, מדובר בהחרפה משמעותית של מגמה קיימת: השימוש הדומיננטי במדדים כמותיים של פרסומים וציטוטים מחרף ביותר

⁷ אכן, ניסיונות לכפות נורמות של דיסציפלינה אחת על דיסציפלינה אחרת התקיימו גם לפני עליית השימוש במדדים אדישי-תוכן, אך השימוש במדדים אלו מחזק מאוד את הנטייה הזו. בפרט, ניסיונות הכפייה בעבר התמקדו ברמת הדיסציפלינה ובאו לידי ביטוי בעימותים פומביים בין עמדות תכניות שונות, ואילו הכפייה של נורמות פרסום (דרך שימוש במדדים אחידים) מתמקדת ברמת החוקרים הבודדים ומתבטאת בדרך כלל בהחלטות לא פומביות אשר מבוססות על הנחות סמויות שאינן עומדות לדיון (ראו בהמשך).

את השינוי במעמד הפרסומים ומהותם. המסר הדומיננטי, אשר עולה מכל עבר, קורא לאופטימיזציה של פרמטרים מסוימים של פרסומים, לא לשיפור איכות המחקר. בפרט, גם כאשר יש התייחסות ל"איכות" הרי שזאת אינה איכות המחקר עצמו אלא איכות זירת הפרסום. מכאן קצרה הדרך להתמקדות בניסיונות לבצע **מניפולציות על המדדים** השונים.

התופעות הללו **מעודדות מחקר קונפורמיסטי** בכיוונים מקובלים ו"פופולריים" שאת תוצאותיו קל יותר לפרסם בזירות "יוקרתיות", ומדכאות **כיווני מחקר חדשניים** ו/או **ביקורתיים** אשר נוטים להיות מודרים מהזירות ה"מרכזיות" ומתפרסמים בזירות שמסומנות כ"שוליות" ע"י עצם העובדה שכיוונים שאינם משתייכים ל"מיינסטרים" זוכים בהם לייצוג משמעותי. האפקטים האלה משפיעים במידה גדולה יותר על מוסדות אקדמיים שחסרים בטחון עצמי ועל תחומי ידע אשר מוחלשים ע"י הניהול האקדמי הנוכחי, אך השפעתם מורגשת גם במוסדות אשר חולשים על הון תרבותי ו/או פיננסי רב וגם בתחומי ידע המועדפים ע"י הניהול האקדמי הנוכחי.

פרק 6 דן בניסיונות **הסנגוריה על השימוש במדדים כמותיים אדישי-תוכן**. נבחין בין סנגוריה אשר טוענת כי השימוש במדדים אלו אינו משנה את מהות תהליך ההערכה לבין סנגוריה אשר תומכת בשימוש במדדים מתוך אמונה שהוא ישנה באופן מהותי ופרוגרסיבי את תהליך ההערכה. לדעתנו, הסנגוריה הראשונה טועה בהערכת יתר של מהימנות המדדים ובאי-אבחנה בעוצמה של הנזקים הישירים והעקיפים, ואילו הסנגוריה השנייה תולה תקוות שווא במה שהיא מזהה כפוטנציאל מקדם-שוויון של המדדים וטועה באי-אבחנה באופי הדכאני שלהם.

פרק 7 נפתח בדיון **בהשתלבות של השימוש במדדים כמותיים אדישי-תוכן בשינוי המטרות של תהליך הערכת האיכות**. מטרתה המקורית של הערכת איכות הייתה לבחון עמידה בתנאי סף מסוימים לשם הקצאה פנימית של משאבים. דוגמאות אופייניות הן החלטה על קידום של חברי סגל אקדמיים וקבלה של מאמרים לזירות פרסום שונות. כמובן שגם בחינות של עמידה בתנאי סף היגן יחסיות, אלא שהמוקד שלהן היא האובייקט הספציפי אשר נבחן ומוערך. בניגוד לכך, המטרה של רוב פעילויות ההערכה המתבצעות כיום הינה **לייצור דירוג והיררכיה**, בעיקר בין יחידות אקדמיות ומוסדות אקדמיים.⁸ יצירת דירוג כזה מחייבת התעלמות מן התכנים השונים והחזונות השונים של היחידות והמוסדות השונים, ומדדים אדישי-תוכן משתלבים בכך באופן מושלם. השאלה שראוי לשאול היא לצורך איזו מטרות מבקשים ליצור דירוג והיררכיה קשיחה ומה התוצאות של יצירתם.

שאלות אלו מתחברות לשאלת הפופולריות של השימוש במדדים כמותיים אדישי-תוכן לצורך הערכת איכות. שאלה זאת מתבקשת לנוכח הביקורת על מהימנות המדדים הללו ולנוכח הנזקים הישירים והעקיפים של השימוש הדומיננטי בהם. ההסברים שיוצגו משתלבים זה בזה ומחזקים זה את זה למרות שמקורם בזוויות הסתכלות שונות.

הסבר שכבר נרמז לעיל, ונראה כחשוב ביותר, הוא השימוש במדדים כמותיים אדישי-תוכן משתלב באופן מושלם **בתרבות הניטור** שהינה נדבך מרכזי בסדר הניאו-ליברלי. המוקד של הסדר הזה הוא החלשת המוסדות החברתיים על ידי הכפפתם לביקורת חיצונית אשר מתעלמת ממטרות החברתית ומתווכ פעילותם של המוסדות ומנתחת אותם במונחים כלכליים גרידא. במקרה של מוסדות אקדמיים, נמחק התוכן המדעי, והניתוח נסוב על תוכן ניהולי ועסקי. העברת המוקד מן התוכן המדעי לתוכן הניהולי והעסקי משתלבת בשינוי יחסי הכוחות בין החוקרים האקדמיים לבין שדרת הניהול האקדמית.

⁸ ניתן לטעון שהיררכיה ואי-שוויון הם **ממהותה** של האקדמיה אלא שטענה זאת מתייחסת להיררכיה אשר נוצרת על בסיס של הישגים מדעיים, ולפיכך בעיקרון זאת היררכיה נזילה אשר משתנה ללא הרף תוך שיקוף של הישגים מדעיים חדשים. ההיררכיה שנוצרת ע"י המדדים נוטה לשמר את עצמה והתלות שלה בהישגים חדשים ובעיקר חדשניים הינה קטנה (מסיבות שנרמזו לעיל).

השימוש במדדים אדישי-תוכן לצורך הערכת איכות אקדמית מאפשר **לבטל את היתרון המובנה של המומחים בידע המדעי הספציפי** ואפילו ליצור יתרון למתמחים בשימוש במדדים הכמותיים (או, אם נדייק יותר, למרגישים בנוח עם שימוש זה). הכפירה בחשיבות התוכן המדעי מערערת את סמכותם של העוסקים בפעילות המדעית (ז"א: העובדים האקדמיים), מבזה את עיסוקם, מעלה על נס את טכניקות הניהול האקדמי, ומספקת לגיטימציה למשטורם של העובדים האקדמיים ע"י שדרת הניהול האקדמית (והמעמד אשר שולט בה מכוח שליטתו בפוליטיקה החיצונית).

המדדים אדישי-התוכן נתפסים כפתרון לקושי של החוקרים האקדמיים עצמם להעריך עבודות מדעיות שמחוץ לתחום התמחותם, עקב רמת הפיצול ההולכת וגדלה בין הדיסציפלינות השונות ולהתמקצעות בתוך כל דיסציפלינה. אלא שהמדדים אינם פותרים את הבעיה אלא מאפשרים, ולמעשה מעודדים, התעלמות ממנה תוך התמכרות לאשליה שהבעיה נפתרה (לפחות מבחינה מעשית).

המדדים הכמותיים אדישי-התוכן גם משתלבים במהלך המודרני של "אובייקטיביזציה" של ידע וייצוגו באופן אשר מסלק את הסובייקט ומרוקן את הידע ממשמעות. למעשה, נכון יותר לומר שהשימוש הדומיננטי במדדים כמותיים אדישי-תוכן של איכות אקדמית מתאפשר באקלים בו התכנים והמשמעות של הפעילות האקדמית מוחלשים ומדלדלים. כמובן ששימוש הפוטנציאל הזה תלוי ביצירה של טכנולוגיה מתאימה, ואכן זאת הופיעה בסוף המאה הקודמת כאשר שימוש במדדים כמותיים (המבוססים על מאגר נתונים ביבליוגרפי) הפך להיות בהישג יד של כל אדם.

ברמה האישית, הסתמכות על מדדים כמותיים מציבה **פיתוי אופורטוניסטי**: היא מאפשרת להחליף את המאמץ האינטלקטואלי הכרוך בקריאה והבנה של חומרים רוויי-משמעות בזריקת מבט חטוף בנתון מספרי בודד. בנוסף, היא מאפשרת להתנער מאחריות להחלטה אשר התקבלה על ידי תלייתה במדדים הכמותיים שהוצגו. כמובן שהמהלך האופורטוניסטי הזה מחייב לגיטימציה או למצער אי-קיום של נורמות אשר מגנות אותו. תנאי זה מתקיים, כמובן, מכוח האקלים החברתי והאינטלקטואלי שתואר לעיל (ז"א תרבות הניטור ופטיש האובייקטיביות, בהתאמה).

המשטר הניאו-ליברלי מבוסס יותר ממשטרים קודמים על שיתוף הפעולה החופשי (מכפיה גלויה) של נתינו. תופעה זאת מסבירה את הצורך ואת הקיום של שליטה עמוקה יותר של המשטר בתודעת הנתינים. מצב זה גם נושא בחובו פוטנציאל של שחרור: המשטר יהיה חסר אונים מול סירוב של חלק מן המערכת לשתף פעולה, ביחוד כאשר מדובר במוסד מרכזי כמו האקדמיה. בפרט, שינוי תהליכי ההערכה באקדמיה וביטול הדומיננטיות של מדדים כמותיים אדישי-תוכן ניתנת לביצוע על ידי האקדמיה עצמה.

2. מדדי איכות אדישי-תוכן

מדדי איכות אדישי-תוכן הם תוצאה של הפעלת כלים סטטיסטיים על מידע ביבליוגרפי הנוגע למאמרים מחקריים, כלומר מדובר בשילוב **מידע בימליומטרי ועיבוד סטטיסטי**.

המידע ביבליוגרפי כולל זיהוי של המאמר, זהות מחבריו, מועד הפרסום, וזירת הפרסום (ז"א כתב-עת, כינוס, מ"ל של ספר, או פרק בספר). מקורו של המידע הביבליוגרפי הוא **בזירת פרסום מוכרת** (להלן: **פרסום מוכר**) או ברשימה הביבליוגרפית המופיעה בסוף מאמר שהתפרסם בזירה כזאת (להלן: **ציטוט מוכר**). מדדים אלו נקראים **אדישי-תוכן** משום שאינם מתייחסים לתוכן המאמרים אלא רק למידע הביבליוגרפי שפורט לעיל.

העיבוד הסטטיסטי מתייחס להתמרה של המידע הביבליוגרפי הנ"ל לכמות יחידה על סקלה מספרית. העיבוד מבוסס על נוסחה מתמטית שלעיתים ידועה למשתמשים במדד ולעיתים לא.

2.1 ריבוי של מדדים

המושג של **זירת פרסום מוכרת** הינו מרכזי לאיסוף מידע ביבליומטרי אלא שהקריטריונים להגדרתו אינם ברורים לרוב המשתמשים בממד מסוים ושונים מממד למדד. בפרט, בחלק מן המדדים הנפוצים, זירות הפרסום המוכרות לצורך הגדרת "פרסום מוכר" אינן זהות לזירות המוכרות לשם הגדרת "ציטוט מוכר". (עובדה זאת כשלעצמה אומרת דרשני.)

בכל מקרה, מדובר בהחלטות מהותיות אשר מתייחסות בעיקר לשתי שאלות: השאלה הראשונה נסובה על **סוג הזירות אשר מוכרות** לצורך הכרה בפרסום ו/או ציטוט (מתוך מגוון אשר כולל כתבי-עת, אוספי מאמרים מסוגים שונים, מו"לים של ספרים, וארכיבים אלקטרוניים מסוגים שונים). ודוק: הרבה מן המדדים מכירים רק בחלק מסוגי הזירות הללו, ואילו מומחים בתחומי ידע שונים מתייחסים באופן שונה לסוגי הזירות השונות.⁹ השאלה השנייה נסובה על **זהות הזירות המוכרות** מבין הזירות אשר מתאימות לכאורה לכל אחד מן הסוגים המוכרים. בפרט, רוב המדדים הקיימים מכירים רק בחלק קטן מזירות הפרסום מן הסוגים המוכרים. כמו כן, תהליך ההוספה של זירות חדשות לרשימת הזירות המוכרות על ידי מדד מסוים אינו מייד, ולעתים עובר זמן רב עד שכתב-עת (או כינוס) חדש "זוכה" להיכלל ברשימה.

בנוסף לקביעת קבוצת הפרסומים והציטוטים המוכרים, להלן **מסד הנתונים הביבליוגרפי**, הגדרת מדד כרוכה **בקביעת סוג העיבוד הסטטיסטי** המופעל על המסד הזה. בפרט, ישנם מדדים המתייחסים רק לפרסומים, ישנם מדדים המתייחסים רק לציטוטים, וישנם מדדים המשלבים פרסומים וציטוטים (בדרכים שונות, ואכן כל שילוב מגדיר מדד אחר). המדדים הפשוטים ביותר הינם מספר הפרסומים או הציטוטים, ואילו מדדים מתוחכמים יותר ממשקלים את הפרסומים על פי זירות הפרסום, דבר המחייב קביעה של משקלים לזירות הפרסום השונות. משקול פופולרי הוא על פי **האימפקט פקטור** של זירת הפרסום, למרות ששימוש זה באימפקט פקטור גרר ביקורת של מפתח האימפקט פקטור עצמו (Garfield 2001). באשר לאימפקט פקטור, הרי שזה מדד אשר מתיימר לקבוע איכות של כתבי עת ע"י השוואה בין מספר הפרסומים בהם למספר הציטוטים של פרסומים אלו.

בדומה לאימפקט פקטור, ישנם מדדים אשר מתיימרים לקבוע את איכות החוקרים על פי נוסחה התלויה במספר הפרסומים ובמספר הציטוטים. המדד הפופולרי ביותר מסוג זה נקרא h -index וערכו מוגדר כמספר הגדול ביותר h של מאמרים שזכו כל אחד ללפחות h ציטוטים. הקוראים יכולים לתמוה, בצדק, מדוע נבחרה הנוסחה הזאת ולא נוסחה אחרת כגון המספר הגדול ביותר h של מאמרים שזכו כל אחד ללפחות $2h$ ציטוטים (או $h/2$ ציטוטים)? התשובה היא שכך החליט פיזיקאי בשם הירש שהציע את המדד הזה (Hirsh 2005).¹⁰

ראוי להבחין בין **מדדים פומביים** ומפורסמים הנמצאים בשימוש רב לבין **מדדים פרטיים** אשר משמשים אדם בודד או יחידה אקדמית בודדת. מדדים מן הסוג הראשון מתיימרים בדרך כלל להיות אוניברסליים (ז"א להתאים לשימוש בכל תחום ידע), בעוד שהמדדים הפרטיים משמשים בדרך כלל להערכה רק באותו תחום ידע עבורו הם פותחו. בפרט, חברי סגל מסוימים אשר סופרים פרסומים של מועמדים בכתבי-עת שהם מחשיבים "יוקרתיים" מפעילים מדד פרטי (שיתכן שמקובל רק עליהם). מדדים פרטיים מסוכנים פחות ממדדים אוניברסליים הן משום שהם מבוססים על מומחיות בתחום הידע בו הם מופעלים

⁹ לרבים מן הקוראים ידוע שמעמד היחסי של ספרים ופרסומים בכתבי-עת שונה בתחומי ידע שונים. מה שידוע פחות הוא שיש תחומי ידע (למשל מדעי המחשב) שבהם פרסום בכנסים מסוימים "נחשב יותר" מאשר פרסום בכל כתב-עת. בתחומים מסוימים (כגון מתמטיקה ופיזיקה) מתייחסים להופעת מאמר בארכיב אלקטרוני כמעט כמו לפרסום בכתב-עת "בלתי יוקרתי" (בניגוד לפרסום בכתב-עת אשר נמנה על "היוקרתיים ביותר").

¹⁰ השרירותיות של הבחירה הזו אינה מנומקת במאמר שלו, אשר מציג שורה של פיתוחים מתמטיים של ההתנהגות הצפויה של המדד במודל מגוון בו חוקרים מפרסמים בקצב קבוע וכל אחד ממאמריהם מצוטט באותו קצב. באופן אופייני, דיון במשמעות מוחלף בדיון במודל מתמטי שרירותי.

והיכרות עם התרבות הפרסומית הנהוגה בו והן משום שהם פתוחים יותר לדיון (מכיוון שהפרטיות הגלויה שלהם מעוררת תהיות לגבי הקביעות העומדות בבסיסן).¹¹ עניין זה יובהר בפרק על ההשלכות הישירות של שימוש במדדים כמותיים.

נקודה שברצוננו להדגיש היא שבבסיס המדדים הכמותיים הנ"ל עומדת שורה ארוכה של החלטות עקרוניות. החלטות אלו מייצגות עמדות בנושאים הנתונים במחלוקת, נושאים הנוגעים למהות הפרויקט המדעי ואופן התנהלותו. יתר על כן, הסבירות והמקובלות של העמדות השונות עשויה להיות תלויה בתחום הידע שעליו מדובר: מה שסביר ומקובל בתחום ידע אחד אינו בהכרח מתאים לתחום ידע אחר. הדיון בעניין זה יורחב בפרק הבא, אבל לפני כן קצת היסטוריה על התפתחות המדדים הנ"ל והתפתחות השימוש בהם.

2.2 קצת היסטוריה

התייחסויות למספר הפרסומים ו/או הציטוטים כמדד לאיכות המחקר האקדמי מופיעות החל מאמצע המאה התשע עשרה, אלא שהתייחסות אלו היו נקודתיות והשפעתן על הקהילה האקדמית עד אמצע המאה העשרים הייתה מצומצמת. בפרט, החל מאמצע המאה התשע-עשרה החלו כמה מ"לים למנות את מספר הציטוטים למאמרים שפורסמו בכתבי-עת שלהם ולהפיץ את המספרים הללו בין מינויים פוטנציאליים כעדות לחשיבות כתב-העת. אך נראה שהשפעת המדד הזה עד לסוף המאה העשרים לא חרגה מעבר להשפעה על מדיניות הרכישה של ספריות אקדמיות.

שימוש במדדים כמותיים אדישי-תוכן להערכת איכות אקדמית משתקף בביטוי "פרסם או התפגר" (publish or perish), שמקורו בשנות השלושים של המאה העשרים. ביטוי זה מייצג התייחסות לעצם הפרסום תוך התעלמות גמורה מתוכנו. סמוך לתקופה זאת גם מופיעה הפרקטיקה של התייחסות למספר הפרסומים הרשומים בקורות החיים של מועמד, ואף לזירת הפרסום שלהם, בדיונים שתכליתם להעריך את איכותו של המועמד. התייחסויות אלו מייצגות שימוש במדד פרטי, אשר מכיל בדרך כלל לתחום הידע הרלוונטי.

איסוף רחב-היקף ועיבוד של מידע ביבליוגרפי ביחס לפרסומים מדעיים החל בשנות השישים (ראו הקמת חברת ISI בשנת 1960). מטרתו המקורית של האיסוף הזה הייתה לשמש תשתית למחקר כמותני של הדינמיקה של תחומי ידע שונים, והשפעתו של מסד הנתונים הזה על הערכת איכות של מחקר אקדמי נשארה מצומצמת בטרם הפכו המדדים הכמותיים אשר מבוססים עליו נגישים לכל.

המחקר של שדות "המדע, הטכנולוגיה והחדשנות" מנקודת מבט כמותית היא הנושא המכונן של תחום ידע הנקרא **סיינטומטריקה** (Scientometrics), אשר נוסד בשנות השישים והשבעים וגובל בסוציולוגיה של המדע ובחקר הספרנות והאינפורמציה (Leydesdorff and Milojevic 2013). נקודות ציון בהתפתחות התחום כוללות את הקמת מסד הנתונים של ISI בשנת 1960 ויצירת מדד SCI בשנת 1964. יסודות עיוניים לתחום הונחו בעבודותיו של Price והתחום מוסד בשנת 1978 עם הופעתו של כתב העת Scientometrics (Price 1978). תחום ידע זה עוסק לא רק בפתוח מדדים כמותיים של פרסומים ובהצדקתם, אלא גם בניסיון להבין את התפתחות של תחומי ידע על ידי ניתוח כמותי של אספקטים שלהם.

¹¹ הפרטיות של המדד מאפשרת דיון בקביעה של רכיביו – כגון כמה חשיבות ניתן ליחס לעצם הפרסום בזירה מסוימת, ומה הפרופיל הנורמלי של פרסומים בתחום – משום שהאדם המשתמש בו נאלץ להכריז על קביעות אלו בעת השימוש במדד (בפרט כאשר השימוש נעשה בנוכחות אנשים אחרים שהמדד לא משותף גם להם).

מסד הנתונים של ISI ומדד SCI נרכשו ע"י החברה המסחרית Thomson בשנת 1992 ונמצאים מאז בבעלותה. בעוד שהחברה המקורית הייתה מחויבת לתמיכה במחקר סיינטומטריקסי, ואפשרה נגישות די חופשית למסד הנתונים שלה, החברה הרוכשת התנערה ממחויבות זאת והגבילה את הגישה למסד עצמו.

החל מתחילת המאה הנוכחית, מידע ביבליוגרפי ומדדים אשר מבוססים עליו הפכו לנגישים בקלות באינטרנט, דבר שהביא לעלייה דרמטית בשימוש במדדים השונים. ההסתמכות הגוברת על מדדים אלו שינתה באופן רדיקאלי את השפעתם, למרות האזהרות והמחאות של הרבה מדענים מן השורה (Adam, 2002).¹²

כמעט כל הפרסומים העוסקים במדדים כמותיים (אדישי-תוכן) של מחקר אקדמי מתייחסים לשאלת המהימנות של המדדים הללו, וכמעט כולם מתייחסים לבעיות ספציפיות במדדים ספציפיים. נקודה נוספת המודגשת בהרבה פרסומים היא שהמדדים משמשים לתכליות שלא להן יועדו.¹³ המאמר הנוכחי הינו ראדיקלי יותר בגישתו: נדונה בו הבעייתיות העקרונית של הרדוקציה של תוכן אקדמי למדד כמותי, וההשלכות המסוכנות של הרדוקציה הזאת.

3. היומרה האובייקטיבית והתלות הסמויה בהחלטות סובייקטיביות

כפי שנרמז בפרק הקודם, בניגוד לאשליה שהמדדים הכמותיים (אדישי-תוכן) של איכות אקדמית הינם "אובייקטיבים" (או "ניטרליים"), הרי שהם מבוססים על שורה של החלטות בלתי-אובייקטיביות שאינן גלויות ולפיכך אינן פתוחות לביקורת. ההחלטה הראשונה והסמויה מכולן היא ההחלטה לקבוע **שוויון בין פרסומים** אשר הופיעו באותה זירה (ובאותה תקופה). אם נתעלם מן ההחלטה הזאת, שהינה עצם ההחלטה לפתח מדד אדישי-תוכן של פרסומים, הרי שההחלטה הראשונה למעשה היא הקביעה **מהו פרסום מדעי**: השאלה היא אלו כתבי-עת, כינוסים, מו"לים וספרים (ואולי אף ארכיבים אלקטרוניים) נחשבים כזירות פרסום אשר נכללים במידע הביבליוגרפי הבסיסי, ושאלה נוספת היא מי מחליט ביחס לשאלה הראשונה ועל בסיס אלו קריטריונים.

בהקשר לשאלה של "מי מחליט" יש להדגיש כי המדדים הפומביים מפותחים, במקרה הטוב, על ידי מומחים לסטטיסטיקה ו/או ביבליומטריקה. אנשים אלו אינם מומחים (ואינם יכולים להיות מומחים) במגוון תחומי הידע אותם המדדים מכסים, ואינם מכירים היטב את התרבות הפרסומית במגוון התחומים הנ"ל. לפיכך, ביחס לרוב תחומי הידע אותם המדד מתיימר לכסות, **ההחלטות של מפתחי המדדים אינן מושכלות ולעיתים קרובות הן מוטעות**: עניין זה יפורט בפרק הבא. נקודה נוספת הראויה להדגשה היא שהרבה מן המדדים מפותחים על ידי חברות מסחריות ו/או מבוססים על מאגר נתונים אשר נבנה על ידי חברות אלו. מיותר לציין ששיקולי חברות אלו הם עסקיים וחפים מכל מחויבות למושג סביר כלשהוא של אמת.

¹² גם ארגונים אקדמיים ממוסדים הביעו את דאגתם מתופעה זאת: ראו למשל סעיף (15) בהכרזת סאן פרנסיסקו ביחס להערכת מחקר (Cagan 2013). הכרזה זאת נחתמה תוך שבועיים ע"י 231 ארגונים אקדמיים, והמוקד שלה הוא התנגדות לשימוש באימפקט פקטור להערכת איכות מחקר. סעיף (15) הינו כללי יותר וקורא לביסוס הערכות איכות על התוכן המדעי ולא על מדדים כמותיים (אדישי-תוכן) של פרסומים.

¹³ אזהרות לגבי מגבלות של המדדים השונים מופיעות גם על גבי אתרים המציעים אותם לשימוש, אלא שרוב המשתמשים לא שמים לב אליהן. לדוגמה, ה"אימפקט פקטור" של כתבי עת פותח במטרה לאמוד את ההשפעה של כתבי עת שונים ולא את איכותם של פרסומים מסוימים בכתבי עת אלו.

לאחר שנקבע מה יחשב כפרסום מדעי, יש להחליט איזה עיבוד סטטיסטי לבצע על המידע הביבליוגרפי הנ"ל. כפי שנאמר בפרק הקודם, קיימים מדדים המתייחסים לפרסומים וקיימים מדדים אשר מתייחסים לציטוטים. מעבר לכך, קיימים מדדים שרק סופרים פרסומים, וקיימים כאלו אשר ממשקלים את המניה ע"י פרמטרים שונים של זירת הפרסום (למשל ה"זיקרתיות" של הזירה – ושוב יש לשאול על אופן קביעתה). קיימים גם מדדים כגון h-index אשר משלבים פרסומים וציטוטים ע"י ערך מספרי יחיד, ובמקרה כזה עולה השאלה של קביעת הנוסחה המתמטית אשר נותנת את הערך המספרי הזה.

ההחלטה איזה עיבוד סטטיסטי (או איזו נוסחה מתמטית) להפעיל אינה החלטה טכנית אלא החלטה אשר משקפת הכרעות עקרוניות לגבי מה מהווה איכות אקדמית. גם אם נניח שהאיכות ממוצה באופן מלא במידע הביבליומטרי, הנחה שאנחנו כופרים בה, הרי שלא ברור כלל מהי הנוסחה אשר מבטאת את המיציף זה: בפרט, איך למשקל את מספר הפרסומים מול מספר הציטוטים? האם לקחת בחשבון רק את מספר הפרסומים והציטוטים של החוקר או האם יש למשקל אותם ע"י מדד כלשהו המתייחס לאיכות זירות הפרסום (ובמקרה כזה איך למשקל ובאיזה מדד של זירות פרסום להשתמש)?

לשם פשטות נניח שהוחלט שלא למשקל את הפרסומים או הציטוטים. במקרה זה נותרנו עם השאלה איך לשלב את מספר הפרסומים ומספר הציטוטים לערך מספרי יחיד. כדי להדגים את השאלה, נתייחס למדען א' שיש לו 20 פרסומים שכל אחד מהם מצוטט 25 פעמים, מדען ב' שיש 15 פרסומים שכל אחד מהם מצוטט 40 פעמים, ומדען ג' שיש לו 50 פרסומים שכל אחד צוטט 12 פעמים. לאיזה מהם יש איכות גבוהה יותר? הנוסחה שבבסיס מדד h-index קובעת למדענים הנ"ל את הערכים 20, 15, 12¹⁴, בהתאמה. אלא שנוסחה שרירותית (לא פחות אך גם) לא יותר הייתה קובעת ערכים של 12, 15, 6, ואילו נוסחה שלישית הייתה קובעת ערכים של 20, 15, 24. נדגיש כי איננו מתייחסים כאן למופרכות של עצם הניסיון לקבוע את איכות המדענים עפ"י האינפורמציה הנתונה לעיל.

ושוב יש לשאול: מי מחליט על אופן העיבוד הסטטיסטי ועל סמך אילו קריטריונים? ולגבי מדדים בהם יש שימוש בפרמטרים של זירת הפרסום יש לשאול איך הם נקבעים: האם ע"י חוות דעת סובייקטיביות (למשל דירוג של זירות פרסום המסתמך על סקר שנעשה בין המומחים "המובילים" בתחום) או ע"י תוצר של סטטיסטיקה של הפרסומים בזירה (למשל מספר הציטוטים הממוצע לפרסום בזירה)? ושוב, איך ומדוע נבחרה דרך קביעה אחת ולא אחרת?

עניין נוסף הראוי להדגשה הוא שבבסיס מדדים כמותיים אדישי-תוכן עומדת נוסחה אחידה להערכת עבודות מחקר על סמך פרמטרים ביבליוגרפיים שלהם. האחידות הזו, שצידה השני הוא סטנדרטיזציה של תהליך ההערכה, מוצגת כיתרון הגדול של הממד, אלא שהיא למעשה חולשתו הגדולה. הממד כופה סטנדרטים ואחידות על תוכן שהוא מגוון באופן יוצא דופן. מדובר, אם כן, בהחלטה שרירותית ובלתי סבירה באופן קיצוני, לכפות על עולם של תופעות מגוונות סרגל סטנדרטי אשר מנוגד למהותו.

נדגיש שוב כי ההחלטות הקובעות את האופי של מדד אדישי-תוכן מסוים הן החלטות סובייקטיביות אשר מבוססות על שיקול דעת, אלא שבדרך כלל זהות המחליטים, ההטיות התרבותיות שלהם, טיב שיקול הדעת שלהם, וההצדקות של ההחלטות שלהם אינן גלויות ואינן פתוחות לדיון. יתר על כן, במקרה של מדדים המתייחסים לפרסומים ביותר מתחום ידע בודד, ההחלטות עשויות שלא להתאים לתרבות הפרסומים בחלק מתחומי הידע אותם הממד מתיימר לכסות. כמוכן, ההחלטות עשויות להיקבע על סמך עמדות ביחס למחלוקות בתחום ידע מסוים, כגון מחלוקות ביחס לעדיפות (שלא לומר בלעדיות) של סוג מסוים של שאלות מחקר ו/או סוג של מתודולוגיות מחקריות, דברים שיש להם השפעה על המעמד של סוג מסוים של פרסומים ו/או זירות פרסום. מקרים אלו מדגישים את הקשר בין ההחלטות הסמויות

¹⁴ כזכור הנוסחה המקורית קובעת את הערך כמספר הגדול ביותר h של מאמרים שזכו כל אחד ללפחות h ציטוטים. הנוסחה השנייה קובעת את הערך כמספר הגדול ביותר h של מאמרים שזכו כל אחד ללפחות 2h ציטוטים. הנוסחה השלישית קובעת את הערך כמספר הגדול ביותר h של מאמרים שזכו כל אחד ללפחות h/2 ציטוטים.

שמזינות את המדדים (אשר נתפסים כ"אובייקטיביים") לבין יחסי הכוחות בין תחומי ידע ובתוך תחומי ידע. ושוב, הבעיה מוחרפת ע"י העובדה שעם החלטות סמויות אי אפשר להתווכח.

3.1 מה מודד המדד?

בנוסף להחלטות הסובייקטיביות אשר נוגעות לאופן העיבוד הסטטיסטי של המידע הביבליוגרפי, הרי שהמידע הביבליוגרפי עצמו הוא תוצאה של החלטות סובייקטיביות: בפרט, ההחלטה **האם לקבל מאמר לפרסום** בזירה מסוימת היא החלטה סובייקטיבית, אשר מבוססת על שיקול דעתם של עורכים ושופטים, וההחלטה **האם לצטט פרסום מסוים** במאמר חדש היא החלטה סובייקטיבית של המחברים. מכאן שגם המידע שעליו מבוססת הסטטיסטיקה הינו תוצר של החלטות סובייקטיביות.

העובדה שמדד מבוסס על החלטות סובייקטיביות של פרטים אינה בעייתית כאשר כל מה שנטען לגבי המדד הוא היותו אגרגציה של החלטות פרטניות, אבל **הטענה שהמדד משקף איכות יחסית של אופציות מסוימות דורשת הנמקה**. לדוגמא: מספר האנשים שרכשו מוצר מסוג מסוים משקף את המספר הזה כשלעצמו בלי קשר לסיבות שהובילו להחלטות הרכישה, אבל היקש ממספר הרכישות לאיכות המוצר דורש הסבר והנמקה. באותו אופן, קישור בין מספר המאמרים של חוקר לבין איכות החוקר (או בין מספר ציטוטים של מאמר לבין איכותו) דורש הסבר והנמקה. מיותר לציין שההסבר המקובל בהקשר של מוצרים מתבסס על השוואה למספר הרכישות של מוצרים "דומים" ושילת האפשרות שפרמטרים אחרים של המציאות משפיעים על העדפות הרכישה. מכאן שההיקש מכמות לאיכות בשדה המוצרים דורש הבנה של השדה אשר חורגת מהערכים הכמותיים. היקש כזה הינו הרבה יותר בעייתי בשדה התרבותי (או האקדמי) בשל השוני העצום בין ה"מוצרים" ובשל הערכים המספריים הקטנים יחסית שאליהם מתייחסים (לדוגמא מספר הפרסומים של מועמד מסוים או מספר הציטוטים של מאמר מסוים).

כדי להמחיש את הטיעון, הבה נדמיין מועמד (בגיל אקדמי מסוים) שיש לו עשרה פרסומים. בהתעלם מן השאלה האם כמות הפרסומים מצביעה על איכות המועמד, נשאל האם כמות זאת הינה קטנה או גדולה? כמובן שהתשובה תלויה בסטטיסטיקה של הפרסומים בתחום הידע הרלוונטי (ובגיל האקדמי הרלוונטי). לצורך עניין זה תחום הידע הרלוונטי אינו ברמה של מחלקה אקדמית, משום שגם ברמת הפרדה זאת ישנה שונות גבוהה בין תתי-תחומים אשר מתבטאת בשיטות מחקר שונות, שאלות מחקר מסוג שונה, ותרבות מדעית שונה (ובפרט שונה בנוגע לאופי הפרסומים).¹⁵ מסתבר שאפילו השאלה "האם עשרה פרסומים זה הרבה או מעט" אינה ניתנת להכרעה ללא מידע נוסף שידוע רק למי שמכיר את תחום הדעת הרלוונטי.

אלא ששאלת כמות הפרסומים רחוקה מלמצות את שאלת האיכות: כך אפילו בעיני הרבה מן המשתמשים במדדים כמותיים לצורך הערכת איכות אקדמית. משתמשים כאלו לא יסתפקו בדרך כלל במספר הפרסומים אלא ישאלו על איכות הזירות שבהם הופיעו הפרסומים הנ"ל. שקלול הפרסומים על פי זירות הפרסום נעשה גם בחלק מן המדדים הכמותיים, אלא שהוא מצריך הערכה של איכות זירת הפרסום ומבוסס על האמונה שאיכות הזירה משקפת את איכות המאמר שפורסם בה. בתנאי הראשון אפשר לטפל רק ע"י שימוש בחוות דעת של מומחים בתחום הידע הרלוונטי ואילו התנאי השני מבוסס על אמונה נאיבית: כל מי שמכיר זירות פרסום אקדמיות יודע שגם השונות בין איכות המאמרים אשר מתפרסמים באותה זירה היא גדולה. בפרט, בזירה הנחשבת ל"יוקרתית" ביותר מתפרסמים, בצד מאמרים פורצי דרך, גם הרבה מאמרים טובים שאינם פורצי דרך, ובפרט מאמרים שאיכותם משתווה לאיכות הרבה מאמרים שמופיעים בזירה פחות "יוקרתית". מכאן, **שמספר המאמרים ושקלולו על פי יוקרת זירות הפרסום השונות, גם כאשר השקלול נעשה ע"י מי שמכירים את התרבות הפרסומית**

¹⁵ גם התייחסות של תחומי הידע השונים הינה פעולה הדורשת הבנה של התוכן של המחקר. ודוק: תחומי ידע שונים אך סמוכים עשויים להיות בעלי תרבות פרסומית שונה ביותר.

של התחום הרלוונטי, אינו יכול לשמש בסיס להחלטה מושכלת בשאלה האם איכות המועמד עומדת בתנאי הסף של הדרישות האקדמיות של היחידה האקדמית. בפרט, האיכות של המאמרים אינה משוקפת באופן מספיק מדויק ע"י היוקרה של הזירות בהן הם התפרסמו.¹⁶

האמור לעיל אינו עומד בסתירה לטענה שיש קורלציה בין איכות המועמד לבין סטטיסטיקת הפרסומים שלו. הטענה היא שמדובר ברמת קורלציה נמוכה מדי עבור החלטה בשאלה האם המועמד עומד בתנאי הסף: רמת הקורלציה יכולה להספיק על מנת להסתכן בהחלטה שלילית במקרים מאוד קיצוניים, אך לא להכרעה לגבי מועמדים סבירים (קל וחומר לגבי הכרעה ביניהם). ברמת קורלציה נמוכה כזאת ולגבי מועמדים סבירים, הכרעה על פי מדדים כמותיים שקולה כמעט להכרעה על פי הטלת מטבע. לדיון נוסף בעניין זה ראו הפרק העוסק בסנגוריה על השימוש במדדים.

דברים דומים תקפים לגבי ניסיון להסתמך על סטטיסטיקה של ציטוטים. גם במקרה זה מספר הציטוטים תלוי מאוד בתרבות הפרסומית של תחום הידע, וגם הציטוטים הינם שונים באיכותם: קיימות מגוון סיבות לציטוט של מאמר מסוים, וזאת אפילו אם נתעלם מן התופעה הנפוצה של ציטוט מאמרים מסיבות בלתי ענייניות (החל מניסיון למצוא חן בעיני המחברים ועד לניסיון לנכס מן החשיבות של כיוון מחקר "נחשב" גם כאשר המאמר המצוטט עוסק למעשה בכיוון אחר)). אמנם הציטוטים הענייניים מצביעים על שימוש ממשי באלמנט תוכני של המאמר המצוטט, אלא שהשימוש עשוי להיות מהותי יותר או פחות למאמר המצוטט, והאלמנט המצוטט עשוי להיות סקירה ביבליוגרפית מקיפה או תרומה טכנית קטנה אך שימושית ולא בהכרח תרומה מהפכנית חשובה.¹⁷ בפרט, פרסום עשוי לזכות להרבה ציטוטים מסיבות שונות הנעות בין היותו פורץ-דרך לבין היותו הסיכום הממצה ביותר של מצב התחום. יתכן שתרומתו היא בהצגת מסגרת מושגית חדשה ויתכן שתרומתו היא בהצגת סימון מוצלח.

3.2 על הטיות הסמויות של המדדים הנפוצים

כפי שהובהר לעיל, מדד כמותי ("אובייקטיבי") מבוסס על שני רכיבים: הרכיב הראשון הוא הגדרת האינפורמציה הרלוונטית ואופן עיבודיה. כאן מדובר בקביעת קריטריונים לגבי 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 מתאם מספק לאיכות (משום שמגוון איכות המאמרים בכל כתב עת הוא עצום).

הרכיב השני של המדד הוא האינפורמציה עצמה. גם כאן, בהקשר של מדדים המתיימרים לשקף איכות אקדמית, מדובר באשליה של אובייקטיביות: הפרסומים והציטוטים אינם אינדיקטורים "אובייקטיביים" של

¹⁶ התופעה הזאת ידועה גם למפתחי מדדים כמותיים אשר ממליצים לא להשתמש באימפקט פקטור של זירת הפרסום כמדד לאיכות של פרסום ספציפי בזירה זו (Garfield 2001). לדוגמא, עיון ברשימת הפרסומים של כל חוקר מוביל מראה שונות גדולה בין איכות מאמרים שהתפרסמו באותה זירה, וכן מאמרים שאיכותם גבוהה יותר אשר התפרסמו כזירות "יוקרתיות" פחות מאשר מאמרים הנופלים מהם באיכותם. מקור השונות הזאת הוא הן בשוני במדיניות הקבלה לפרסום של עורכים שונים והן בשוני במדיניות ההגשה לפרסום של מחברים שונים. בפרט, קיימים עדיין מחברים ישרי-דרך אשר בוחרים בזירת הפרסום על פי התאמתה למאמרם ולא על פי יוקרתה. בפרט, יתכן שהמחברים סבורים שהמאמרם מתאים יותר לכתב-עת של תת-תחום בו התפרסמו מאמרים באותו נושא ו/או שפרסום בכתב-עת זה יעשה אותו נגיש יותר לקהל קוראים רלוונטי יותר, למרות שהוא ראוי להתקבל לכתב-עת כללי יותר אשר נחשב "יוקרתי" יותר. יתכן גם שהמחברים מאמינים בחשיבות הפוטנציאלית של המאמר, אבל מעריכים שבשלב זה הוא יישפט רק על סמך התוצאות הקונקרטיות שהוא משיג בפועל.

¹⁷ הציטוטים במאמר זה מספקים דוגמה טובה לטענה זאת.

איכות (שאותה מבקשים לאמוד), אלא תוצר של החלטות סובייקטיביות (סמויות). שאלת הפרסום של מאמר בזירת מסוימת מוכרעת בדרך כלל על סמך חוות הדעת של מספר קטן (בד"כ שנים או שלושה) של שופטים אנונימיים, שאיש פרט לעורכים אינו יודע את טיב שיקול הדעת (הסובייקטיבי) שלהם. לא רק שמדובר בגורם סובייקטיבי, אלא שלמשתמש במדד אין מידע על מידת מומחיותו ומהימנותו של גורם זה. ואילו לגבי ציטוטים, הרי שגם כאן מדובר בהחלטות סובייקטיביות (של המצטטים) 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 ביחס לשונות (הסטטיסטית) של הפרטים?¹⁸ ומה מידת הקורלציה הסטטיסטית המושגת?

אולם החשש המרכזי מתייחס לקביעת קריטריונים. המדדים של פרסומים וציטוטים מתייחסים לכתבי-עת מסוימים ולעתים כוללים שקלול שלהם. מי קובע את רשימת כתבי-העת הנלקחים בחשבון ומי קובע את אופן שקלולם ועל סמך אילו שיקולים? לפחות בכל הנוגע לתחום 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.¹⁹ כמו כן, 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, בשל הטיות תרבותיות של מפתחי המדדים.

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 בהרבה מקרים אינן אלא תופעות "מערביות" ו/או "אנגלוסקסיות"). ההכרעה בשאלה של קריטריונים למדדי פרסומים כרוכה כאן בשאלה העקרונית של תכני המחקר, ומכאן הסכנה שבהפקרתה למפתחי מדדים כמותיים: שאלות עקרוניות לגבי תכני מחקר צריכות להיות מוכרעות בדיון פתוח של הקהילה המדעית הרלוונטית.

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?)

בנוסף לכך, כפי שנאמר בתחילת הפרק הנוכחי, הרבה מן המדדים מפותחים על ידי חברות מסחריות ו/או מבוססים על מאגר נתונים אשר נבנה על ידי חברות אלו. השיקולים של חברות אלו הם עסקיים, והן נוטות שלא לחשוף את הקריטריונים שהנחו אותם בפיתוח המדדים שלהם. השיקולים העסקיים יכולים להכתיב העדפה של קלות המימוש של המדד על מהימנותו, שלא לציין את האפשרות של הטיית המדד על מנת לשרת אינטרסים עסקיים של מו"לים מסחריים. לנוכח עובדות אלו, משונה עוד יותר להתייחס לערכי המדדים הללו כאילו היו שיקוף ניטרלי של איכות אקדמית.

מן האמור לעיל נובע שהתבססות על מדדים אדישי-תוכן להערכת תרומה מחקרית – שהינה דבר של תוכן – היא דבר תמוה ובלתי-מושכל, על פי הגדרתו. בפרקים הבאים נראה שזהו גם דבר מזיק ביותר.

¹⁸ במקרה של פרסומים מדובר בשונות הגבוהה של מהימנות חוות הדעת ביחס למספר הקטן יחסית של חוות דעת. במקרה של ציטוטים השונות מתייחסת גם לסיבות השונות לציטוט.

¹⁹ לדיון בהחלטות הבעייתיות אשר עומדות בבסיס האימפקט פקטור ראו סקירתם של (Glanzel and Moed 2004). לביקורת על השימוש באימפקט פקטור להערכת איכות פרסומים ראו (Seglen 1997). לתיאור של מניפולציות שיחיות על האיפקט פקטור ראו סקירתם של (Falagas and Alexiou 2008).

כפי שיפורט בהמשך, הנזק החמור ביותר הוא עקיף, אלא שראוי לפתוח בנזק הישיר שהוא קבלת החלטות תת-אופטימלית.

4. השלכות ישירות של השימוש במדדים כמותיים אדישי-תוכן

החלטות הנוגעות לשכירה וקידום של חברי סגל הינן בעלות חשיבות מכרעת לאקדמיה, שהרי עמוד התווך של האקדמיה הוא חברי הסגל: איכות האקדמיה כקהילה של חוקרים הנמצאים בחזית הידע בתחומם נקבעת על ידי איכותם של אותם חוקרים. לפיכך, כל ירידה באיכות קבלת ההחלטות בנושאים אלו הינה פגיעה באיכות האקדמיה. בפרק זה נטען כי השימוש במדדים כמותיים אדישי-תוכן להערכת איכות של מחקר אקדמי פוגעת באיכות ההחלטות המתקבלות. נתחיל בדיון בחוסר המהימנות של סוגים ספציפיים של מדדים כמותיים ומשם נעבור לדיון כללי יותר בחוסר המהימנות המהותי של מדדים כמותיים להערכת איכות מחקר.

4.1 חוסר המהימנות של סוגים ספציפיים של מדדים אדישי-תוכן

נבחין בין כמה פרמטרים ו/או עקרונות העומדים בבסיס המדדים הנפוצים.

- **ספירת פרסומים:** כפי שנאמר בפרק הקודם, המגוון של עבודות המחקר הוא עצום גם אם נתעלם מעבודות פורצות-דרך ומ"פרסומי זבל". לקוראים המבקשים להיווכח בשוני העצום שבין האיכות של מחקרים שונים מוצע להתבונן ברשימת הפרסומים של מדען (סביר) כלשהו בתחום המומחיות שלהם.
- משקול פרסומים על פי **זירת הפרסום:** שוב, מגוון הפרסומים בכל זירת פרסום עצום, וכדי להיווכח בו די לעיין במאמרים שפורסמו בזירת פרסום (סבירה) כלשהי בתחום המומחיות של הקורא. בנוסף, בהרבה תחומי ידע איכות זירת הפרסום אינה משתקפת באימפקט פקטור שלה.
- ספירת ציטוטים ו/או משקול פרסומים עפ"י **מספר הציטוטים:** גם כאן המגוון עצום. כדי להיווכח בכך מוצע לקוראים לבחור מאמר סביר כלשהו בתחום המומחיות שלהם ולעקוב אחר מגוון הציטוטים השונות לציטוט של פרסומים במאמר זה.

עדויות קונקרטיות לחוסר המהימנות של מדדים ספציפיים באה לידי ביטוי באי היציבות שלהם ובכך בנטייה שלהם שלא להסכים אחד עם השני. לדוגמא, **הדירוג של אוניברסיטאות מסוימות עפ"י המדדים המקובלים משתנה משנה לשנה באופן תזזיתי למדי**, בעוד שסביר להניח שהאיכות של כל אוניברסיטה היא דבר יציב יחסית. מספר הפרסומים והציטוטים לפי מדדים שונים (למשל Google Scholar מול Web of Science) נותן ערכים שונים, בעיקר בשל השוני בהגדרת זירות הפרסום המוכרות ע"י מדדים שונים. בנוסף, יש לציין שהאינפורמציה במסדי הנתונים הנמצאים בבסיס המדדים השונים מכילה שגיאות רבות, בעיקר בזיהוי שמות המחברים (חלק מן הטעויות הן בפרסום המקורי וחלקן נובעות מטעות בשליפת המידע מן הפרסום המקורי), והאמונה שהשגיאות מתקזזות היא חסרת בסיס (כי ברור שהן מושפעות מגורמים כמו מורכבות שם החוקר ומידת פרסומו).

נגדיש שאין באמור לעיל בכדי לטעון שאין כל קורלציה בין הפרמטרים הנ"ל לבין איכות המדען המפרסם. טענתנו היא שהקורלציה חלשה מדי והיא מובהקת בעיקר במקרים קיצוניים, בעוד שמה שרלוונטי לשימוש הנפוץ במדדים הוא המהימנות שלהם לצורך הערכת מקרים רגילים. טיעון זה יורחב בפרק בו נדון בסוגריה על השימוש במדדים כמותיים.

4.2 חוסר המהימנות המהותי של כל מדד אדישי-תוכן

הקשר בין המדדים הכמותיים השונים לבין איכות אקדמית לא נבדק **ברצינות** מעולם. בדיקה רצינית הייתה מחייבת מדגם גדול של פרסומים בכל אחד מתחומי הידע המכוסים על ידי המדד, איסוף חוות דעת

מקצועיות על איכות הפרסומים האלו והשוואתם לערך המספרי הניתן על ידי המדד. השאלה הדורשת ברור אינה אם קיימת קורלציה כלשהי בין האיכות למדד, אשר באה לידי ביטוי בעיקר במקרים קיצוניים (כגון עבודות מופת או להבדיל "פרסומי זבל"), אלא האם הקורלציה גבוהה מספיק על מנת להצדיק הסתמכות עליה במקרים לא קיצוניים (ובפרט, לצורך הערכת איכות של כמעט כל העבודות המדעיות ו/או כמעט כל חברי הסגל האקדמי). ההנחה הנאיבית שהקורלציה בין איכות העבודות לבין ערכי המדד גבוהה מספיק (לשם הסתמכות על המדד) היא ספקולטיבית.²⁰ 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.

הטענה שמדד מסוים משקף איכות מסוימת היא טענה המתייחסת לתוכן מסוים, ובחינתה היא עניין למומחים לתוכן זה. פיתוח ובחינה של מדד המשקף מחלות מסוימות הוא עניין למומחים לרפואה, 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 להבחין על ידי מומחים לתחום זה, ולא על ידי אנשים שאינם מבינים את התכנים המקצועיים של אותו תחום וגם עלולים שלא להכיר ולהבין את התרבות האקדמית שלו.

הצורה היחידה להעריך, באופן מהימן, איכות של מחקר אקדמי 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.

יש לציין שגם מומחים בתחום ידע מסוים לא יצליחו ליצור מדד אדיש-תוכן עם מתאם גבוה (במידה מספקת) לאיכות אקדמית בתחומם. הסיבה היא שאפילו המחקר בתחום ידע מצומצם ביותר הוא רב-גוני מכדי שאפשר יהיה לעשות לו רדוקציה למדד מספרי מן הסוג המדובר. בפרט, כפי שכבר נאמר ובניגוד 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 – שונות באיכותן (אשר בלתי אפשרי לפתח נוסחה 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 לגבי מתן קביעות לחברי סגל אקדמי ו/או טענה שמחלקה אחת טובה יותר ממחלקה אחרת, וכך הלאה. אני מבקש להעיד על כך שלמרות מומחיותי בתחום המחקר שלי והיותי בעל כישורים מתמטיים וניסיון רב בבניית מודלים, אינני יכול להציע מדד כזה לתחום המומחיות שלי.

²⁰ אפשר לטעון שגם הטענה הספקנית לפיה הקורלציה נמוכה (מכדי להצדיק הסתמכות כזאת) הינה ספקולטיבית. אלא שבמקרה כזה ראוי להעדיף את הטענה הספקנית (ראו הנמקה בסוף פרק זה). הטענה הספקנית שמדובר בקורלציה נמוכה מבוססת על הרושם של הרבה מומחים בתחומי ידע שונים, ובפרט זה הרושם שלי. רושם זה מבוסס על הניסיון שלי בהערכה של מחקר אקדמי, בעיקר בתחום ההתמחות שלי, על ההבנה שלי את תחום הידע הזה, ועל ההיכרות שלי עם המדדים הכמותיים השונים. כאמור, הספקנות כלפי מהימנות המדדים השונים נתמכת על ידי מספר רב של מדענים: ראו למשל דיווח של מנהל רשות הערכת המחקר הבריטית על פיה הרשות לא משתמשת במדדים משום 96% מן המדענים שנשאלו על כך העדיפו חוות דעת של מומחים (Ball 2005).

בשלב זה אנו עומדים בפני שתי טענות מנוגדות שלא נבדקו באופן יסודי מספיק: טענה **נאיבית** שמניחה שהמדדים הכמותיים הקיימים, או למצער חלקם, משקפים איכות אקדמית באופן מספק להסתמכות מכרעת עליהם, וכנגדה טענה **ספקנית** אשר גורסת שהמתאם בין המדדים הכמותיים ואיכות אקדמית נמוך מכדי להצדיק הסתמכות כזאת.²¹ מכיוון שהמדדים הכמותיים מתיימרים לשקף דבר אחר אשר בו אנו מעוניינים, הרי שבמקרה של ספק ממשי לגבי איכות השיקוף, ראוי שלא להשתמש בו (ז"א לא להפעיל קיצור דרך שמהימנותו נתונה בספק).

לסיכום פרק זה, אציין שלא רק ש"המדדים הכמותיים (אדישי-התוכן) לאיכות אקדמית" אינם מהימנים, אלא **שאשליית האובייקטיביות** שהם מייצרים הינה מסוכנת ביותר. ראשית, משום ש"האובייקטיביות" הזאת **מסתירה עמדות עקרוניות הנתונות במחלוקת** ביחס לתכנים של תחומי ידע. מחלוקות אלו הינן לגיטימיות, ויותר מכך הן מפרות. אך חשוב שהן תהנה גלויות ושניתן יהיה לדון בהן במקום ובזמן המתאים ולצמצם את השפעתן במקומות אחרים (בפרט על ידי גילוי מלא של השפעת המחלוקת העקרונית על הערכת האיכות האקדמית הספציפית). סכנה שנייה בדומיננטיות של המדדים הכמותיים אדישי-התוכן הינה שהיא מעודדות **הזנחה של הערוץ המהימן** היחיד להערכת איכות אקדמית:²² חוות דעת המבוססות על הבנה ושיקול דעת.

5. השלכות עקיפות של השימוש במדדים כמותיים אדישי-תוכן

כפי שנרמז לעיל, הנזקים העקיפים של השימוש הדומיננטי במדדים כמותיים אדישי-תוכן להערכת איכות חמורים עוד יותר מן הנזקים הישירים. הנזק הברור ביותר הוא בדחיפה נוספת של האקדמיה ממחויבות למחקר כנה ואמיץ, שסימנו המובהק הוא האיכות, להתעסקות בכמות של פרסומים. כמובן שמטרת המחקר היא לא רק לגלות ידע חדש או צורת הסתכלות חדשה אלא גם לשתף את הקהילה המדעית הרלוונטית בידע הזה, והשיתוף הזה נעשה דרך המדיום של פרסומים. עובדה זאת לא משנה את הראיה של המחקר כמטרה ואת הפרסום כאמצעי לשיתוף. הדומיננטיות של המדדים הכמותיים אשר אדישים לתוכן מערערת את הראיה הזאת: באווירה הנוצרת, **הפרסומים הופכים למטרה, ואילו המחקר נחווה כאמצעי להשגתם.**

החלפת התפקידים בין אמצעים למטרות משתלבת בריקון הפרויקט המדעי ממשמעות: התפיסה של המחקר האינדיבידואלי כחלק מפרויקט קהילתי וכחלק משיחה משמעותית עם קהילת החוקרים

²¹ הסתמכות שולית על מדדים כמותיים, תוך ראייתם ככלי עזר גס, הינה בעייתית פחות. התומכים בשימוש במדדים כמותיים טוענים ששימוש זהיר (ובפרט, בלתי-מכריע) במדד בעל מהימנות נמוכה עשוי להועיל, אלא שהם מתעלמים מן הנזקים העקיפים אשר נגרמים ע"י השימוש הזה אשר יפורטו בפרק הבא. בנוסף לכך, התומכים ממעיטים בהערכת הנזקים הישירים אשר נוצרים כאשר לא קיימת מודעות למהימנות הנמוכה של המדדים ויותר מכך כאשר מעדיפים את "עדוּתם" של המדדים על חוות דעת של מומחים לתוכן הרלוונטי אשר מתייחסים לתוכן הזה. לגבי העדפת עדותם של המדדים ראו גם הערה 2222. לדיון נוסף בסנגוריה על שימוש במדדים, ראו פרק בנושא זה.

²² הצד הישיר והמיידני יותר של ההזנחה הזאת מתבטאת בהעדפת "עדוּתם" של המדדים על חוות דעת של מומחים לתוכן הרלוונטי אשר מתייחסים לתוכן הזה. בטווח ארוך יותר ובאופן עקיף יותר, מוביל התהליך לניוון של חוות הדעת המקצועיות (וירידה באיכותן) שנגרמת כאשר הסוקרים מעדיפים להביא ערכים של מדדים כמותיים אדישי-תוכן במקום להתייחס לתוכן עצמו. מיותר לציין שהתופעה השנייה נובעת מן הראשונה הן משום שחלק מן הסוקרים מעדיפים גם הם מדדים כמותיים אדישי-תוכן על חוות דעת והן משום שחלק מן הסוקרים מניחים שקוראי חוות הדעת נותנים אמון רב יותר במדדים כאלה. חשוב להדגיש שלא מדובר כאן בסכנה פוטנציאלית אלא בתופעה ממשית שגלויה לעינין של כל המצוי בסוד תהליכי הקידום באקדמיה.

הרלוונטיים מוחלפת בחוויה של מרדף חסר משמעות אחר פרסומים ו/או תחרות חסרת משמעות בין החוקרים השונים (שהרי לתחרות אין כל משמעות מעבר לזכייה בה).

הפיכת הפרסומים האקדמיים מאמצעי למטרה אינה חידוש של המדדים הכמותיים, אלא תהליך היסטורי שמלווה את הפיכת המחקר האקדמי מתחביב אריסטוקרטי למקצוע (יוקרתי). בשלב הראשון של התהליך, ניתן היה להבחין בעליית ההתייחסות למספר פרסומים בהערכה של חברי סגל אקדמיים: תחילה הייתה ההתייחסות למספר הפרסומים מבוישת במקצת, ואילו כיום מספר הפרסומים משמש כנימוק גלוי ומכריע בהרבה דיונים של ועדות הערכה. כמו כן, בפגישות רשמיות ושיחות חברותיות מועבר מסר לאקדמיים המתחילים לפיו כדאי להם להתרכז בהגדלת מספר פרסומיהם. בשלב השני, שהחל לפני כשני עשורים, ניתן להבחין בהסתמכות גוברת על מדדים כמותיים מתוחכמים יותר (ובפרט סטטיסטיקה של ציטוטים). כפועל יוצא מכך, ציון מספר הציטוטים של כל מאמר ברשימת הפרסומים של מועמד הינה נורמה בקהילות מדעיות מסוימות. המסר העולה מן הנורמה הזאת היא שיש להתרכז לא רק בהגדלת מספר פרסומים אלא בפרמטרים נוספים כמו מספר הציטוטים והיוקרתיות של זירות הפרסום.

לעיתים קרובות נטען כי המסר האחרון הוא מסר הקורא לשיפור איכות המחקר, אלא שטענה זאת שגויה. המסר קורא לאופטימיזציה של פרמטרים מסוימים של פרסומים, לא לשיפור איכות המחקר. בפרט, לא נאמר "כדאי לכם לבצע מחקר שישיע באופן משמעותי על תמונת העולם של מומחים בתחום המחקר שלכם" אלא "כדאי לכם לפרסם בכתב-עת יוקרתי יותר ולזכות למספר רב יותר של ציטוטים." ודוק: ברמה המילולית, לא מדובר על איכות המחקר כשלעצמו, אלא על האפקט שלו על מדדים כמותיים. ומהלכה למעשה: הניסוח הנ"ל מזמין התרכזות בערך של המדד ובניסיון לשפר אותו בלי שיפור של איכות המחקר.

ההתרכזות בערך של המדד היא הזמנה למניפולציות עליו. ראשית, מדובר בתימרוץ של ניסיונות "לדחוף" מאמרים לזירות פרסום שאינן מתאימות להם, משום שה"רווח" במקרה של קבלה הוא גדול ואילו הנזקים לקהילה המדעית אינם נספרים. (הנזקים לקהילה כוללים פרסומים בזירות בלתי מתאימות, בזבוז זמנם של שופטים, ועיכוב בפרסום).²³ שנית, מדובר בתימרוץ של חוקרים לבקש מחוקרים אחרים לצטט את מאמריהם של הראשונים או ביצירת "בריתות" של ציטוט הדדי. מיותר לציין, שלכל המניפולציות הללו אין שום קשר לקידום איכות המחקר, נהפוך הוא: השתרשות של התנהגות בלתי-עניינית היא מכשול בפני קידום איכות המחקר.

יתר על כן, הדגש על פרסום בזירות "יוקרתיות" ועל מספר ציטוטים (בטווח של שנים בודדות) מעודד מחקר קונפורמי אשר מתאים לפרדיגמות המקובלות ומתיישר עם האופנות הדומיננטיות. כל חריגה מן המקובל מסתכנת בירידה בפופולריות אשר מתבטאת בירידה בערך המדדים הנ"ל. סקרנות אינטלקטואלית ופתיחות מחשבתית, שלא לציין ביקורתיות של ממש, מדוכאות משום שיש להם מתאם שלילי עם פופולריות. כל זה פועל בכיוון של שימור התכנים המקובלים ודיכוי של חדשנות רעיונית.

כיווני מחקר חדשניים, גישות חדשניות ובעיקר ביקורתיות, ותחומי ידע חדשים אשר מצליחים להתקיים למרות הדיכוי הנ"ל, מודרים מן הזירות ה"יוקרתיות" ומתפרסמים בזירות חדשות המוקמות לצורך כך, אלא שזירות אלו מסומנות על ידי הממסד האקדמי כ"בלתי יוקרתיות". מדדים אשר מבוססים על סולמות יוקרה האלו מעניקים למאמרים האלו ערך נמוך יותר (ולעתים לא מחשיבים אותם כלל).²⁴ כך משוכפלים

²³ בהנחה שעדיין יש לפרסומים חשיבות החורגת מתביעת בעלות על טקסט, הרי שיש חשיבות לבחירת זירת הפרסום: הזירות המתאימות הן אלו הנקראות על ידי המומחים בתחום הידע הרלוונטי. פרסום בזירות אחרות מסתכן בהחמצת הקוראים הללו ומהווה בזבוז של המשאבים של הזירות האחרות. במקרה של דחיה מזירת פרסום בלתי מתאימה בעליל, מדובר בבזבוז של משאבי שיפוט.

²⁴ בפרט, כפי שצוין בפרק "מדדי איכות אדישי-תונן", רוב המדדים הקיימים מכירים רק בחלק מזירות הפרסום (מן הסוגים המוכרים). כמו כן, תהליך ההוספה של זירות חדשות לרשימת הזירות המוכרות על ידי מדד מסוים אינו מידי, ולעתים עובר זמן רב עד שכתב-עת (או כינוס) חדש "זוכה" להיכלל ברשימה. עובדות אלו, בשילוב עם

יחסי הכוחות בין ה"מרכז" האקדמי וה"פריפריה" שלו בחסות של מדד "אובייקטיבי" שהינו אדיש-תוכן: האי-שוויון נשמר ע"י תהליך סמוי מן העין שבו כל מחקר בתחום המוחלש מקבל ערך קטן יותר (על פי המדדים הנ"ל) מערך של מחקר דומה לו אשר משתייך ל"מיינסטרים" האקדמי.

בהקשר של מדינה פריפריאלית, כגון ישראל, ראוי להדגיש גם את **האפליה כנגד מחקרים העוסקים בתרבות המקומית**, משום שקשה לפרסם מחקרים כאלו בזירות הבין-לאומיות המרכזיות, ואילו פרסומם בכתבי-עת מקומיים נושא ערך נמוך יותר על פי המדדים השונים (אם הוא נחשב בכלל). בפרט, הרבה מן המדדים אינם מחשיבים כלל פרסום בכתבי-עת בעברית, ופרסומים אלו זוכים לפחות ציטוטים.

לסיכום, השימוש הדומיננטי במדדים כמותיים אדיש-תוכן (בעיקר פומביים וכלליים) מדכא שלושה סוגים של מחקר: עבודות פורצות-דרך מדוכאות הן ע"י הדגש על כמות הפרסומים והן ע"י הדגש על ציטוטים (בטווח זמן קצר יחסית). עבודות ביקורתיות מדוכאות ע"י הדרתם מן הזירות המרכזיות ומתן ערך נמוך לזירות הפרסום בהן עבודות אלו מופיעות. עבודות העוסקות בתרבות המקומית מדוכאות באופן דומה. דוק: תהליך הדיכוי הינו אוטומטי, בלתי נראה, ואינו פתוח לדיון.

לסיום פרק זה, ברצוני לחזור ולציין כי מדידה אדיש-תוכן של מחקר מדעי **פוגמת אנושות בתחושת היעוד המדעי**: תחושת השייכות לקהילה אשר עוסקת בפרויקט מרתק אובדת ומוחלפת בתחושה של בדידות, תלישות ומלחמת קיום. הפגיעה הזאת, אשר **מסכנת את מהות הפרויקט האקדמי**, נובעת מן האווירה שנוצרת כאשר חלקים דומיננטיים במערכת מבצעים רדוקציה של תוכן המחקר לניקוד בתחרות חסרת משמעות.

6. הסנגוריה של המשמשים במדדים הכמותיים אדיש-תוכן

יש מקום להבחין בין סנגוריה (שתקרא להלן **שמרנית**) אשר מבקשת להגן על השימוש במדדים תוך טענה ששימוש זה משרת את תהליך ההערכה הקיים לבין סנגוריה (שתקרא להלן **מהפכנית**) אשר רואה את השימוש במדדים כאמצעי לשינוי משמעותי בתהליך ההערכה הקיים.

6.1 הסנגוריה השמרנית

התומכים ("השמרנים") בשימוש במדדים כמותיים אדיש-תוכן טוענים ששימוש זהיר (ובפרט, בלתי-מכריע) בהם עשוי רק להועיל גם כאשר מהימנותם נמוכה,²⁵ אלא שהם מתעלמים הן מן הנזקים העקיפים אשר נגרמים ע"י השימוש הזה (אשר פורטו בפרק 5 (שכותרתו "השלכות עקיפות של שימוש במדדים")) והן מן הנטייה של המדדים הללו להפוך לדומיננטיים בתהליך ההערכה בשל החזות האובייקטיבית שלהם (ראו סוף פרק 4 (שכותרתו "השלכות ישירות של שימוש במדדים")). התעלמות זו, כשלעצמה, חוטאת בניתוח חלקי של הבעיה. אלא שגם נימוקי הסנגוריה כשלעצמם הינם שגויים. נימוקים אלו מסתמכים על הטענה שקיימת **קורלציה חיובית בין המדדים לאיכות אקדמית** ועל הטענה הגנרית **שאינפורמציה נוספת ודאי שאינה יכולה להזיק**. להלן ניתוח שתי הטענות.

בעניין **הקורלציה חיובית בין המדדים לאיכות אקדמית**, הרי שהיא גלויה לעין במקרים קיצוניים: נדיר למצוא (בימינו) מדען מוכר ומוערך שאין לו אף פרסום, קשה למצוא מאמר חשוב שלא זוכה למספר רב

העובדה שבמצב הנוכחי השתייכות לרשימת הזירות המוכרות ע"י המדדים השונים חשובה לרוב האקדמאים, מוסיפה על הקושי בהקמת זירת פרסום חדשה. הקושי הזה עומד בדרכם של כיווני מחקר חדשים וגם בדרכם של יוזמות להקמת זירות פרסום עם "גישה חופשית" (Open Access), תופעה אשר משחקת לידי המו"לים המסחריים.

²⁵ לנוכח הספקות שהובעו בפרקים הקודמים, קשה להתייחס ברצינות לטענה שמהימנות המדדים הכמותיים גבוהה.

של ציטוטים (בהנחה שעבר "מספיק זמן" מאז פרסומו), וקשה למצוא פרסום "חסר ערך" שזוכה להמון ציטוטים. אפשר לנסות להרחיב את הטיעון הזה למקרים פחות קיצוניים, אלא שהמתאם בין איכות לכמות נהפך לפחות מובהק ככל שמתרחקים מן הקצה.²⁶ הבעיה היא שהשימוש הרווח במדד אינו במקרי הקצה אלא במקרים הרגילים, ודווקא לגבי תחום זה לא ברור שיש בכלל מתאם בין איכות לכמות. יתר על כן, רמת המתאם אשר מתקיימת במקרי הקצה מיוחסת (ולו באופן מוחלש – אבל לא מוחלש דיו) גם למקרים הרגילים. מכאן, שבשימוש הרגיל במדד מיוחסת לו מהימנות גבוהה מזאת שיש לו למעשה – מהימנות שנרכשה במקרי-קצה ולא הותאמה מספיק למקרים הרגילים. מסתבר שגם אם יש למדד מתאם חיובי עם איכות אקדמית במקרים שאינם מקרי-קצה, הרי שהשימוש במדד במקרים אלו מייחס לו מהימנות הגבוהה בהרבה ממהימנותו בפועל.

הטענה שאינפורמציה נוספת ודאי שאינה יכולה להזיק נשמעת טאוטולוגית, אלא שהיא אינה נכונה בהרבה סיטואציות אנושיות ובפרט אינה נכונה בהקשר הנוכחי. המהימנות הגבוהה אשר מיוחסת ללא הצדקה למדדים הכמותיים גורמת להזנחה של מקורות מידע אחרים (קרי: חוות דעת מקצועיות) ולניוונם של המקורות הללו. התוצאה המיידית היא העדפה של מקור אינפורמציה מהימן פחות על מקור אינפורמציה מהימן יותר (או למצער הרעשת מקור האינפורמציה המהימן יותר על ידי המקור המהימן פחות), ואילו בטווח הארוך נפגעת האיכות של המקור המהימן יותר.

אכן, שורש הבעיות הוא בנטייה לייחס למדדים הכמותיים מהימנות העולה בהרבה על מהימנותם בפועל. מקור הנטייה הזאת הוא במראית העין האובייקטיבית של המדדים, אשר נובעת מעצם היותם מדדים כמותיים. כמובן שהנזק עולה ככל שמיוחסת למדדים מהימנות רבה יותר, ובפרט כאשר לא קיימת מודעות למהימנות הנמוכה של המדדים האלו.

6.2 הסנגוריה המהפכנית

הביקורת על תת-הייצוג הבולט באקדמיה של **קבוצות אוכלוסייה מוחלשות** מוביל חלק מן המבקרים לשים את ייבם על מדדים כמותיים אדיש-תוכן כאמצעי לביטול תת-הייצוג הנ"ל. ההצעה הזאת מניחה כמה הנחות שגויות. ראשית, היא מזהה את המקור העיקרי של תת-הייצוג בתהליכי ההערכה של חוקרים, בעוד שהמקור העיקרי של תת-הייצוג הוא בהסללה של אוכלוסיות מוחלשות ללימודים מוחלשים ובחוסר פיצוי על שוליות חברתית בשלבים הרבה יותר נמוכים של מערכת החינוך. שנית, המציעים מניחים ששימוש במדדים מנטרל הטיות תרבותיות, בעוד שחלק מן ההטיות הללו בניית לתוך המדדים עצמם ובכל מקרה המדדים לא מפקיעים את ההחלטה מן הוועדות הרלוונטיות. שלישית, בהנחה הסבירה שיש מתאם חיובי בין מוצא סוציו-אקונומי נמוך לבין עניין במחקר ביקורתי, מקומי ו/או פורץ-דרך, הרי שהעובדה שהמדדים מדכאים מחקרים מסוגים אלו אומרת שהם פועלים בממוצע כנגד חוקרים ממוצא זה.²⁷

²⁶ לדוגמה, בדיקת עשרים המאמרים המצוטטים ביותר שלי (על פי Google Scholar) מגלה מתאם נמוך בין מספר הציטוטים לבין איכות המאמרים (על פי חוות דעתי המקצועית): חמשת העבודות החשובות ביותר שלי אכן ממוקמות בעשיריה הראשונה (עפ"י סדר מספר הציטוטים), אלא שהן לא מופיעות בסדר חשיבותן, וביניהן מופיעות עבודות חשובות הרבה פחות. לגבי חמש עשרה העבודות האחרות אינני רואה כל מתאם בין איכותם היחסית לבין מספר הציטוטים היחסי להן הן זוכות. המצב הרבה יותר חמור ביחס לספירה על פי Web of Science: על פי ספירתם, העשיריה הראשונה כוללת רק שתיים מבין חמש העבודות החשובות ביותר שלי.

²⁷ הניתוח הנ"ל תקף, אך במידה פחותה יותר, לגבי תת-הייצוג המגדרי: גם כאן יש אלמנטים של הסללה (בדמות עידוד חלש יותר של יצירתיות ובולטות), הטיה תרבותית, ומתאם חיובי של מגדר עם עניין במחקר חברתי ביקורתי.

ההנחה השנייה מתקשרת לרצונם המבורך של התומכים ("מהפכנים") להשתמש במדדים כמותיים אדישי-תוכן על מנת לקעקע היררכיות של כוח באקדמיה. הנחתם היא שהשליטה של "חברי המועדון" בתהליכי השכירה והקידום תחולש אם תחוזק ההשפעה של מדדים כמותיים אדישי-תוכן על קבלת ההחלטות. העובדה שכמעט כל "חברי המועדון" תומכים בשימוש במדדים כמותיים צריכה לשמש תמרוך אזהרה: אכן, בניגוד לאמונתם של "המצדדים המהפכנים", שימוש במדדים כמותיים פתוח יותר למניפולציות, שהרי תמיד אפשר לבחור באיזה מדד להשתמש, ופסק דינו של המדד פתוח הרבה פחות לדין וערעור.²⁸

גרסה מתונה יותר ומסתברת יותר של ההנחה השנייה טוענת ששימוש במדדים כמותיים אדישי-תוכן עשוי לנטרל הטיות תרבותיות בלתי מודעות ובלתי מכוונות. אכן, ההנחה שהטיות תרבותיות בלתי-מודעות פועלות כנגד אוכלוסיות שסובלות מתת-ייצוג באקדמיה אינה בלתי סבירה, למרות שנראה לי שמגזימים בהשפעתה. בכל מקרה לא נראה לי נכון להתמודד עימה על ידי החלפת הניסיון להערכה מושכלת ע"י שימוש במדדים שמשמעותם אינה ברורה, קל וחומר שהתועלת העשויה לצמוח מכך, אם בכלל, הינה קטנה ביחס לנזקים שנסקרו לעיל.

7. הסברים לפופולריות של מדדים כמותיים אדישי-תוכן

לנוכח הספקות בדבר מהימנותם של מדדים כמותיים אדישי-תוכן להערכת איכות אקדמית, ראוי לתהות על ההסתמכות הנרחבת על מדדים אלו שנפוצה גם בקרב אנשי אקדמיה. בפרט, התהייה היא לגבי שימוש במדדים אלו כפוסקים מרכזיים ואף אחרונים בשאלות של איכות אקדמית. לפני שנדון בשאלת הפופולריות של המדדים בקרב החוקרים הפעילים באקדמיה, ראוי לדון בפופולריות שלהם בבירוקרטיה האקדמית והממשלתית שהיא אשר מקדמת בנחישות את השימוש במדדים אדישי-תוכן.

7.1 המעבר מהערכה איכות פנימית לדירוג חיצוני

המטרה המקורית של הערכת איכות אקדמית היא לבדוק עמידה בתנאי סף. אמנם גם תנאי הסף הם דבר יחסי, משום שאלו נקבעים על סמך הערכה של איכות של כלל המחקר האקדמי בתחום הידע הרלוונטי ו/או התחומים הסמוכים לו, אלא שההערכה מתמקדת באותה עבודה (או קבוצת עבודות) שאותה מעריכים. לעומת זאת, דירוג של עבודות אקדמיות הוא פעילות שמטרתה קביעת יחס סדר בין עבודות אקדמיות: המטרה אינה לבדוק עמידה בתנאי סף מסוימים אלא לדרג.

דירוג של עבודות אקדמיות או של חוקרים אקדמיים ניתן להצדקה במקרים שיש הכרח להקצות משאבים מוגבלים. לדוגמא, במקרה שבכנס מסוים ניתן להציג מספר מוגבל מראש של עבודות או במקרה שיחידה אקדמית יכולה לשכור רק מועמד אחד מבין מספר מועמדים אשר כולם עומדים בתנאי הסף. אלא שבעשורים האחרונים משמשים דירוגים של עבודות, חוקרים ויחידות אקדמיות למטרות אחרות שאין דבר בין לבין קידום האיכות האקדמית. העיסוק האובססיבי בדירוג של חוקרים ויחידות אקדמיות מסתמן כאמצעי שליטה של הבירוקרטיה האוניברסיטאית והממשלתית, ולמצער כאימוץ בלתי-ביקורתי של המיתוס העסקי של תחרות. (ודוק: התחרות העסקית היא על דירוג, בעוד שמה שמכונה "התחרות האקדמית על ממצאים מדעיים" הינו חתירה להשגת הממצאים עצמם ולא לדירוג בתחרות.)

²⁸ הערה זאת מניחה שהוועדות משתמשות במספר מדדים כמותיים ומפעילות אותם על פי שיקול דעתן ולא על סמך נוסחה קבועה מראש. במקרה כזה מרווח הפעולה אינו יכול להתמלא בהתייחסות לתוכן אשר לא נמצא במדדים עצמם, ומכיוון שהבחירה אינה יכולה להתבסס על התוכן הענייני הרי היא מופקרת לשרירות לב. בנוסף, יש להביא בחשבון שדווקא עבודות ביקורתיות נפגעות ופגעו משימוש במדדים כמותיים אדישי-תוכן אשר מבוססים על הטיה לטובת המקובל (קרי: המקובל על "חברי המועדון" שקרובים יותר חברתית למפתחי המדדים).

הדירוג הוא במהותו תהליך חיצוני לתוכן: מטרתו לדרג תכנים אשר שונים במהותם, ובדרך כלל ענינו אינו בערך הדירוג הניתן לתוכן מסוים אלא ביצירת דירוג של התכנים. הדירוג משמש לרוב גורמים חיצוניים למערכת ובפרט לצורכי החלטות על תקצובה. האידאולוגיה השלטת גורסת שאיכות המערכת תשתפר על ידי תחרות שתזן על ידי טיפוח הטובים יותר באמצעות תקצוב נדיב יותר והענשת הטובים פחות על ידי צמצום תקצובם. אלא שמדיניות זו גורמת להחלשת החלשים יותר, שאינם נעלמים (שהרי מישהו תמיד יתפוס את תחתית הדירוג), ופוגעת בכלום משום שהיא מסיטה את תשומת הלב מהמטרות המקוריות (קרי: איכות אקדמית) למיקום בדירוג. המדיניות הזו מטריפה את כל המערכת על ידי כפייה של השתתפות במשחק סכום אפס שבו כל הצלחה של אחד היא כישלון של השני. תחרות זו מחבלת בשיתופי הפעולה והסולידריות בין המוסדות והחוקרים, ומגדילה את חוסר הוודאות שמצדו מפריע לתכנון ארוך טווח.

ההתמקדות בדירוג יחסי של מחקר אקדמי משתלבת היטב באופי הנומרי של המדדים הכמותיים. מדדים כמותיים מגדירים דירוג באופן (כמעט) מיידי, בעוד שקביעת דירוג מלא על סמך חוות דעת מילוליות (אשר מתייחסות לתכנים) היא משימה קשה ביותר. כמובן שאין פירוש הדבר שהמדדים הינם יותר אינפורמטיביים ביחס לאיכות המחקר, אלא שהם יוצרים אשליה של עושר אינפורמטיבי על ידי יצירת בלבול מושגי בין הייצוג המדויק של מידע סטטיסטי ובין התופעה שהמידע הזה מייצג במקור (אם בכלל).²⁹

7.2 ההשתלבות של המדדים הכמותיים בתהליכים חברתיים ותרבותיים כלליים

היומרה האובייקטיבית של המדדים הכמותיים אדישי-התוכן משתלבת היטב באקלים אינטלקטואלי אשר אידאל הידע שלו הוא מכניסטי ושואף למזער את השפעת הסובייקט על המחקר המדעי (Daston and Galison 2007). באקלים כזה, מראית העין של אי-תלות בסובייקטים מפתה ביותר. זאת ועוד, גירוש הסובייקט מתחולל בד בבד עם גירוש המשמעות, זאת אומרת עם הניסיון להימנע מדיון במשמעות של דברים ולהסתפק במעקב אחר ערכים נומריים המיוצרים באופן מכני. כמובן, שהמדדים הכמותיים אדישי-התוכן מתאימים לכך.

היומרה האובייקטיבית גם משתלבת היטב באווירה חברתית של אי-אמון וחשדנות כלפי כל מוסד ציבורי וכלפי הסוכנים המקצועיים של המוסדות הנ"ל (Shore 2008).³⁰ באווירה שכזאת כל חוות-דעת מקצועית חשודה בפניות ונחשבת פסולה לעדות, ומקודמת פנטזיה על מנגנון הערכה חיצוני, נייטלי ובלתי ניתן להטיה. המדדים הכמותיים אדישי-התוכן נחווים כמממשים את הפנטזיה הזאת משום שהם **מסתירים** היטב את החלטות העקרוניות אשר עליהן הם מתבססים, את ההטיות התרבותיות בהן הם נגועים, ואת האינטרסים אותם הם משרתים.

²⁹ הייצוג המדויק יותר של אותן עובדות יוצר אשליה של אמינות גבוהה יותר: נניח שנעשו 81 ניסיונות וב-49 מתוכם נצפתה תופעה מסוימת. דיווח על הופעת התופעה ב-61% מן המקרים נשמע אמין יותר מאשר דיווח על הופעתה ב-60% מן המקרים. זאת למרות שהיחס 49/81 שווה בקירוב ל-0.6049 ולמרות שיחס זה לא מייצג את המציאות עצמה אלא רק את הניסויים שנעשו.

³⁰ מעניין לציין שאי-אמון וחשדנות מופנים בעיקר כלפי מוסדות ציבוריים ועובדיהם, בעוד שחברות עסקיות ומנהליהן נהנים בדרך כלל מאמון יחסי. מצב זה הוא תוצר של האידאולוגיה הנאו-ליברלית, אשר קוראת לחיזוק הפיקוח החיצוני על מוסדות ציבוריים והחלשת הפיקוח החיצוני על חברות עסקיות והחלפתו בפיקוח עצמי (Shamir 2008). בנוסף לאירוניה בכך שחושדים במי שמטרתו היא שירות הציבור ונותנים אמון במי שמטרתו היא שירות עצמו, יש אירוניה בכך שמתיימרים למדוד כמותית דווקא שירותים ציבוריים (כגון חינוך, בריאות, ורווחה) שמהותם הינה איכות (ובפרט את איכות האקדמיה).

בפרט, המדדים הכמותיים אדישי-תוכן מתאימים באופן מושלם ל**תרבות הניטור** (Audit Culture) שהינה חלק מרכזי מן ההגמוניה הניאו-ליברלית (Shore 2008). מנקודת מבט זו מסתבר שיומרתם של המדדים לתאר את המציאות הינה כיסוי למהותם כאמצעים לשינוי המציאות: שימוש במדדים כמותיים כבסיס להקצאת משאבים גורם למערכות האקדמיות לכוון את פעילותן באופנים אשר מגדילים את ערך המדד שלהן. כאשר המדד מוגדר באופן התואם את הסדר הקיים ו/או הרצוי, התוצאה היא דחיפה אוטומטית וסמויה של האקדמיה לשירות של הסדר הזה.³¹ בפרט, השימוש במדדים הנפוצים כופה על המערכת האקדמית שינוי של מטרות **משירות ארוך טווח של החברה** על ידי גילוי של ידע חדש והעשרת בני החברה בתכנים אקדמיים **לשירות המעמד החברתי הדומיננטי והשלטון הנוכחי** על ידי פיתוחים יישומיים והכשרת "כוח אדם" מקצועי. לפיכך, המדדים הכמותיים הם אמצעים למשטור האקדמיה ע"י המשטר הניאו-ליברלי תוך שיתוף פעולה (לא בהכרח מודע) של שדרת הניהול האקדמי.³²

בנוסף לכך, השימוש במדדים כמותיים אדישי-תוכן של איכות אקדמית **מקודם על ידי המעמד הברוקרטי האקדמי** ("המנהלים") כאמצעי שליטה שלהם במעמד החוקרים בפועל ("העובדים"), משום שהשימוש במדדים מבטל את היתרון המובנה של החוקרים בפועל המתבטא בהבנה של התוכן הרלוונטי. שימוש במדד כמותי אדישי-תוכן לא דורש שום הבנה בתוכן האקדמי, והתדמית האובייקטיבית של המדד מאפשרת לברוקרט האקדמי (אשר המתייחס למדד ברצינות תהומית) להופיע כמקצוען האמיתי (משום שהמומחה האקדמי מתקשה להתייחס ברצינות למדד). בנוסף לכך, הדגש על המדד אדישי-תוכן מכחיש את חשיבות התוכן של המחקר עצמו ולמעשה מתעלם מן התוכן אשר מעסיק את החוקר הפעיל. מכיוון שתוכן זה מכונן את זהותו של החוקר הפעיל, הרי שמדובר בדיכוי של החוקר הפעיל. חיזוק מעמדו של הברוקרט האקדמי והחלשת החוקר הפעיל מסתכמים ב**שינוי יחסי הכוחות בין המעמד הברוקרטי של האקדמיה לשאר העובדים האקדמיים**.³³

הדומיננטיות של הסתמכות על מדדים כמותיים אדישי-תוכן משתלבת בטרנספורמציה של האקדמיה ממוסד מסוג מיוחד שלא מקדם מטרות שנקבעות מראש ומבחוץ למוסד ציבורי רגיל שמטרותיו נקבעות מראש ומבחוץ. באקלים שבו כל המוסדות הציבוריים מנהלים כאילו היו חברות עסקיות, משמעות הדבר היא שגם האקדמיה צריכה להתנהל כחברה עסקית. הטרנספורמציה הזאת מצדיקה דיפוזיה של הגישה העסקית (והאידיאולוגיה הניאו-ליברלית) אל השדה האקדמי. שאלות שהובנו בעבר כשאלות אקדמיות מובנות כיום כשאלות ארגוניות. בפרט, הצגת בעיית הערכת האיכות של מחקר אקדמי כבעיה ארגונית (ולא כבעיה אקדמית) מצדיקה את השימוש בכלי ניהול סטנדרטיים לטיפול בה. בפרט, הסטנדרטיזציה הקיצונית אשר באה לידי ביטוי במדדים הכמותיים, נהפכת לנורמה מחייבת, על עף שהיא מנגדת וזרה לרוח האקדמיה.

פרספקטיבה מערכתית אחרונה מתייחסת לתבנית של השתלטות הטכנולוגיה על התוכן כסימפטום של השתלטות הראציונליות האינסטרומנטלית על הראציונליות המהותית. במקרה זה משתלטת הטכנולוגיה הביבליומטרית על התכנים המדעיים, תוך העלמתם מן העין.

³¹ מנגנוני הרגולציה המדינתיים פועלים באופן אוטומטי וסמוי בכיוון של כפיית מטרות ועקרונות המוצגים כניטרליים, על ידי ניטור המבוסס על מדדים אשר תואמים את המטרות והעקרונות האלו. אלא שכפי שהובהר בפרקים הקודמים, המטרות והעקרונות הללו אינם ניטרליים. בפרט, הם שונים מן המטרות והעקרונות אשר הנחו את האקדמיה בעבר.

³² חשיבות שיתוף הפעולה הזה, מנקודת מבט רגולטורית, מודגשת במאמרו של קינג (King 2006).

³³ השכבה של הברוקרטיה האקדמית כוללת את כל ראשי היחידות האקדמיות וכן בעלי תפקידים נוספים ביחידות אלו, כולל את אלו הנמצאים במעבר בין תפקיד לתפקיד. להערכתי מדובר ביותר מחמשה אחוזים של הסגל האקדמי.

7.3 הסברים נוספים

כפי שנטען בתת-הפרק הקודם, ההשתלבות של המדדים הכמותיים בתהליכים חברתיים ותרבותיים כלליים מהווה הסבר ראשי לקבלת השימוש בהם כלגיטימי ואפילו כמועדר.

הסבר נוסף נוגע לתהליכי השינוי אשר עוברים על האקדמיה עצמה. בפרט, הכוונה לפיצול ההולך וגדל בין הדיסציפלינות השונות ולהתמקצעות בתוך כל דיסציפלינה אשר יוצרים מציאות בה אפילו חברי אותה יחידה אקדמית מתקשים להבין את מחקרי עמיתיהם, קל וחומר להעריך אותם.³⁴ המדדים אדישי-התוכן מוצעים כתרופת פלא לקושי המהותי העולה מן המציאות הזו, אלא שלא מדובר בתרופה אלא בסם הזיה אשר מעלים את הבעיה מן התודעה אך לא מן המציאות. בניגוד לכך, פתרון מושכל לבעיה מחייב הכרה בכך שאין ברירה אלא להסתמך על חוות דעת של אחרים, אשר מסוגלים להעריך את העבודות, ויש להתרכז בזיהוי מומחים שניתן לסמוך על חוות דעתם (משימה אפשרית אך לא קלה כשלעצמה).

שינוי נוסף אשר מתרחש באקדמיה הוא עליה בכוחה ומעמדה של שדרת הניהול האקדמית. שינויים אלו בולטים במיוחד בישראל, ובעיקר בעקבות דו"ח ועדת מלץ (2000). כפי שצוין, השימוש במדדים כמותיים משרת את השליטה של שדרת הניהול האקדמית בחוקרים בפועל וכן את משטור האקדמיה כולה על ידי הכוח הפוליטי החיצוני.

אם נעבור מהסברים מערכתיים להסברים פסיכולוגיים, הרי שהשימוש במדדים כמותיים מונע מעצלות מחשבתית ומרצון להתנער מאחריות: הסתמכות על מדדים כמותיים מייצרת את הניסיון להתמודד עם תכנים מורכבים וטיעונים מסובכים אשר מופיעים בעבודה האקדמית עצמה ו/או בחוות הדעת של מומחים בתחום הרלוונטי. המידע הכמותי זמין באופן מיידי ואילו קריאה ביקורתית של חוות דעת (שלא לדבר על מאמרים מקוריים) היא עניין המצריך זמן ומאמץ מחשבתי. בנוסף לכך, הסתמכות על מדד כמותי מאפשרת להתנער מאחריות להחלטה, משום שההסתמכות על המדד לא נחוות כהחלטה (למשל: "היית רוצה לתמוך בקידום המועמד, אבל מספר הפרסומים/ציטוטים קטן מדי").

ההסבר הפסיכולוגי הנ"ל מצביע על אינטרס אופורטוניסטי של חברי אקדמיה, אינטרס פרטי ומיידי אשר עומד בסתירה לאינטרסים הכלליים וארוכי-הטווח שלהם. אינטרס אופורטוניסטי זה מתלכד עם תהליכים חברתיים ותרבותיים כלליים, ועם אינטרסים פוליטיים של המשטר הניאו-ליברלי ושל הבירוקרטים האקדמיים אשר משתפים עמו פעולה. מדובר, באוסף של תהליכים, כוחות ואינטרסים שפועלים כולם באותו כיוון ולפיכך מחזקים זה את זה.

מסקנות

הדבר היחיד שניתן לומר לזכותם של המדדים הכמותיים אדישי-התוכן הוא שמדדים אלו מצויים בשפע וקלים לשימוש. כמובן שזאת גם סיבה מכרעת 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.

³⁴ הפיצול הדיסציפלינרי המדובר כאן הינו רב יותר מן הפיצול שבין "שתי התרבויות" או "שלושת התרבויות".

לשימוש הגובר במדדים כמותיים אדישי-תוכן יש מספר השלכות עקיפות הרסניות. אלו כוללות את הזנחת התוכן והמשמעות ואת התמקדות בטכניקה ובמכניות, את שכפול יחסי הכוחות והגדלת האי-שוויון בתוך תחומי ידע וביניהם, ואת דיכוי תחושת הייעוד והסקרנות האינטלקטואליות. בנוסף לכך, אותם הלכי רוח אשר תומכים בהסתמכות הגוברת על מדדים כמותיים, גם מתחזקים על ידי השימוש הדומיננטי במדדים אלו. מדובר בחיזוק ההתמכרות למכניזם והתעלמות ממשמעות, בהתעצמות האופורטוניזם (עצלנות מחשבתית והתנערות מאחריות), בחיזוק השליטה של המעמד הברוקרטי ("המנהלים") במעמד החוקרים בפועל ("העובדים"), תוך הגדלת דיכויים של האחרונים, ובחיזוק ההשפעה של הגישה העסקית והאידיאולוגיה הניאו-ליברלית על השדה האקדמי.

בנוסף, יש עודף התעסקות בנושא הערכת האיכות האקדמית והזנחה של פעולות אשר באמת ממוקדות בשיפור האיכות האקדמית. התעסקות היתר בהערכת איכות, ובפרט תוך שימוש במדדים כמותיים אדישי-תוכן, מהווה אמצעי שליטה ומשטור של האקדמיה תוך שינוי מטרתיה. בניגוד לכך, נראה שיש תמימי-דרך אשר מאמינים שעצם המדידה בצירוף לחץ על חברי האקדמיה לשפר את המדדים שלהם תוביל לשיפור האיכות. לדעתי, זו טעות: שיפור האיכות של האקדמיה מחייבת הגדלת המשאבים העומדים לרשות כל חלקיה, בניגוד למדיניות הנוכחית שמשמעותה האובייקטיבית היא יבוש תחומי ידע נרחבים. בנוסף, יש צורך ברענון של האתוס האקדמי ולא בדיכוי.

השימוש במדדים כמותיים אדישי-תוכן לצורך הערכת איכות אקדמית מותנה בהסכמת האקדמיה לפרקטיקה הזאת ובשיתוף פעולה עמה. התנגדות נחושה של רוב גדול של החוקרים לפרקטיקה הזאת לא תאפשר לה להתקיים: שדרת הניהול של האקדמיה והכוחות הפוליטיים החיצוניים אינם יכולים לכפות פרקטיות אקדמיות ואינם יכולים לדכא מרידות רחבות היקף באקדמיה. האפקטיביות הפוטנציאלית של התנגדות "מן השטח" גדולה כאן יותר מבכל מקום אחר.

תודות

מאמר זה חב את קיומו ללין חלזין-דברת: לין היא שדחפה אותי להפוך את הביקורת העקרונית שלי (על מדדים כמותיים אדישי-תוכן) למאמר מפורט, ששאלה אותי שאלות מפתח שהובילו אותי לכמה וכמה אבחנות, ושהציעה לי רעיונות מעניינים משלה והפניות קריאה מועילות. בנוסף ללין, ברצוני להודות לדקלה בייטנר, ארי ברטל, הילה דיין, גילי דרורי, תמר הגר, אייל חוברס, שי לביא, דוד לוי-פאור, מנחם (מני) מאוטנר, יצחק (יאני) נבו, אהוד פינס, הגר קוטף, ויוסי שוורץ על דיונים מעניינים שהרבה ממה שנאמר בהם השפיע על המאמר.

מקורות

מלץ, דו"ח (2000). דין וחשבון של הוועדה הציבורית לבחינת המבנה הארגוני של המוסדות להשכלה גבוהה בראשותו של שופט ביהמ"ש העליון (בדימוס) יעקב מלץ.

מקינטאיר, אלסדייר (2006). מעבר למידה הטובה. הוצאת שלם.

סאלברג, פאסי (2015). לומדים מפינלנד: תובנות ממערכת חינוכית מוצלחת. הקיבוץ המאוחד; מכון מופ"ת; מכללת סמינר הקיבוצים; הסתדרות המורים.

Adam, David (2002). Citation analysis: the counting house. *Nature*, 415, 726-729.

Ball, Philip (2005). Index aims for fair ranking of scientists. *Nature*, 436, 900.

- Cagan, Ross (2013). San Francisco Declaration on Research Assessment. *Disease Models & Mechanisms*. Editorial, 6.
- Daston, Lorraine, and Peter Galison (2007). *Objectivity*. Zone Books, New York.
- Espeland, Wendy Nelson, and Sauder, Micael (2007). Rankings and reactivity: how public measures recreate social worlds. *American journal of sociology*, 113 (1), 1-40.
- Falagas, Matthew E., and Vangelis G. Alexiou (2008). The top-ten in journal impact factor manipulation. *Archivum Immunologiae et Therapiae Experimentalis*, 56 (4), 223-226.
- Garfield, Eugene (2001). Impact factors, and why they won't go away. *Nature*, 411, 522.
- Glanzel, Wolfgang, and Moed, Henk (2004). Journal impact measures in bibliometric research. *Scientometrics*, 53 (2), 171-193.
- Halffman, Willem, and Radder, Hans (2015). The academic manifesto: from an occupied to a public university. *Minerva*, Springer, DOI 10.1007/s11024-015-9270-9, Brief Communication.
- Hirsch, Jorge (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*, 102 (46), 16569-16572.
- Laloë, Franck, and Mosseri, Remy (2009). Bibliometric evaluation of individual researchers: not even right... not even wrong! *Europhysics News*, Vol. 40, No. 5, 26-29.
- King, Roger (2006). Analysing the higher education regulatory state. *Discussion paper 38, centre for the analysis of risk and regulation*. London: London School of Economics.
- Leydesdorff, Loet, and Milojevic, Stasa (2013). Scientometrics. In arXiv:1208.4566. Forthcoming in: Lynch, M. (editor), *International Encyclopedia of Social and Behavioral Sciences* subsection 85030.
- Price, D.J. De Solla (1978). Editorial statement. *Scientometrics*, 1 (1).
- Seglen, Per O. (1997). Why the impact factor of journals should not be used for evaluating research. *BMJ*, 314, 498-502.
- Shamir, Ronen (2008). The age of responsabilization: on market-embedded morality. *Economy and society*, 37(1), 1-19.
- Shore, Cris (2008). Audit culture and illiberal governance (universities and the politics of accountability). *Anthropological Theory*, 8 (3), 278-29.
- Snow, Charles Percy (1959). *The Two Cultures and the Scientific Revolution*. Cambridge University Press.