

הערכת עמיתים מקוונת והיכרות בין לומדים: האם הם תלויים זה בזה?

מירי ברק

מאיה אושר

הטכניון – מכון טכנולוגי לישראל

bmiriam@technion.ac.ilushermaya@tx.technion.ac.il

Online Peer Assessment and Students' Familiarity: Are they Interdependent?

Maya Usher

Miri Barak

Technion – Israel Institute of Technology

Abstract

Peer assessment is a process in which learners evaluate the work of their peers. As the number of online distance learning courses increase, peer assessment is becoming more popular since it provides an efficient way for grading numerous open assignments when the teaching resources are limited. Yet, the quality of peer assessment in diverse learning environments has not received sufficient attention. Hence, this study examined the quality of online peer grading and feedback in three learning environments: Classroom (on-campus course), SPOC (small private online course), and MOOC (massive open online course). The Classroom group ($n = 50$) were university students, highly familiar with each other. The SPOC group ($n = 50$) were university students who took the course online and were somewhat familiar with each other. The MOOC group ($n = 50$) were people from different countries who were strangers to each other. All the students were required to submit the same learning assignments and to provide online peer assessment. Findings indicated that the assessment quality of the Classroom students was higher than that of the other groups. The grades awarded by them were in higher correlation to those given by the teaching assistants, and the feedback that they have provided was better associated with higher order thinking levels.

Keywords: Familiarity, Higher education, Massive open online course, Online peer assessment, Online peer feedback.

תקציר

הערכת עמיתים הינה תהליך בו לומדים שופטים את איכות העבודה של עמיתיהם. ככל שגדל מספר הקורסים המקוונים הניתנים מרחוק, כך נעשה שימוש רב יותר בהערכת עמיתים, שכן זו מהווה דרך יעילה להערכת מספר רב של משימות פתוחות בעוד משאבי ההוראה נותרים מוגבלים. עם זאת, בחינת האיכות של הערכת עמיתים מקוונת בסביבות למידה מגוונות לא זכתה לתשומת לב מחקרית רבה. לפיכך, מחקר זה בחן את איכות ציוני ומשובי העמיתים בשלוש סביבות למידה: כיתת לימוד (קורס שניתן בקמפוס), SPOC (קורס מקוון קטן ופרטי) ו-MOOC (קורס מקוון פתוח רב משתתפים). קבוצת הכיתה ($n = 50$) כללה סטודנטים הרשומים לאוניברסיטה, בעלי היכרות גבוהה זה עם זה. קבוצת ה-SPOC ($n = 50$) כללה סטודנטים מהאוניברסיטה אשר נרשמו לקורס בגרסתו המקוונת, בעלי היכרות מסוימת שכן הם למדו יחדיו בקורסים קודמים. קבוצת ה-MOOC ($n = 50$) כללה אנשים ממדינות שונות שהיו זרים זה לזה. כל הסטודנטים התבקשו להגיש את אותן המשימות

הלימודיות ולספק הערכת עמיתים מקוונת. הממצאים הצביעו כי איכות ההערכה בקרב קבוצת הכיתה הייתה הגבוהה ביותר בהשוואה לשאר הקבוצות. הציונים שהסטודנטים מקבוצת הכיתה העניקו לעמיתיהם היו במתאם גבוה יותר עם ציוני המתרגלות והמשובים המילוליים שנתנו לעמיתיהם הצביעו על חשיבה מסדר גבוה.

מילות מפתח: היכרות בין סטודנטים, הערכת עמיתים מקוונת, השכלה גבוהה, משוב עמיתים מקוון, קורסים מקוונים פתוחים מרובי משתתפים.

מבוא

הערכת עמיתים הינה אסטרטגיית הוראה מוכרת ומוערכת בה הלומדים לוקחים חלק בשיפוט והערכת תוצרי הלמידה של עמיתיהם (Sadler & Good, 2006; Topping, 1998). גוף הולך ומתרחב של מחקר מדגיש את ערכה הרב של הערכת העמיתים לתהליך הלמידה (Barak & Rafaeli, 2004; Falchikov & Goldfinch, 2000). היא תועדה כמעלה מוטיבציה בקרב הלומדים (Hanrahan & Isaacs, 2001), כמקדמת את תהליך הלמידה (Li & Gao, 2016; Price et al., 2013) וכמגבירה את רמת המעורבות של הלומדים (Bloxham & West 2004).

על אף הפוטנציאל הרב שיש לתהליך הערכת עמיתים לקידום הלמידה, הערכת עמיתים תועדה כצורכת זמן רב וכזו שאינה נלקחת מספיק ברצינות מצד הלומדים (Hanrahan & Isaacs 2001; Snowball & Mostert, 2013). באשר להטיות על בסיס היכרות מוקדמת בין הלומדים, מחקרים דיווחו על ממצאים נוגדים. חלקם מצאו כי התהליך נתפס בעיני תלמידים אחדים כתחרות על פופולאריות ולכן נתון להטיות על בסיס חברויות המתבטאות בהענקת ציונים גבוהים לחברים (Falchikov, 1995). מנגד, חוקרים אחרים לא מצאו השפעה של חברות והכרות מוקדמת על גובה הציונים בתהליך הערכת עמיתים (Azarnoosh, 2013; Love, 1981; Morahan-Martin, 1996).

אחד ההיבטים במרכז ההתעניינות של אנשי חינוך וחוקרים בתחום הערכת עמיתים, הינו מידת האמינות והדיוק של הציונים הניתנים על ידי סטודנטים בהשוואה לאלו הניתנים על ידי צוות ההוראה (Liu & Carless 2006; McConlogue, 2015). במחקרים מסוימים נמצאו מתאמים חיוביים בין הציונים שהוענקו על ידי עמיתים לאלו שניתנו על ידי המורה (Sadler & Good, 2006; Topping, 1998) או המרצה (Kearney, Perkins & Kenndy-Clark, 2016; Li, et al., 2016); אך מחקרים אחרים מצאו שסטודנטים נטו לתת לעמיתיהם ציונים גבוהים יותר (Kulkarni et al., 2013; Sitthiworachart, 2016). על פי Sunol ועמיתיו (2016), ניתן לשייך ממצא זה לחוסר הניסיון של הסטודנטים במתן הערכה או לתחושת סולידאריות עם עמיתיהם.

ישנם גורמים שעשויים להשפיע על מידת הדיוק של ציוני העמיתים ועל קרבתם לאלו הניתנים על ידי צוות ההוראה. דוגמאות לכך הן אנונימיות (Falchikov 1995; Falchikov & Goldfinch, 2000), הצגת מחוון ברור ומדויק ומתן הדרכה לביצוע הערכה נכונה (Harris, 2011; Kearney et al., 2016). משתנה נוסף אשר נמצא כמשפיע על מידת ההלימה בין ציוני העמיתים לציוני המתרגלות הוא הוספה של משוב עמיתים, משמע, הערות כתובות המסייעות בשיפור העבודה (Falchikov & Goldfinch, 2000; Li, et al., 2016; Suen, 2014). במחקרים בהם בוצע תהליך של משוב עמיתים נמצא כי התהליך מפתח תוצרי למידה ברמה גבוהה (Barak & Rafaeli, 2004), מעודד חשיבה ביקורתית (Liu & Carless, 2006; Nicol, Thomson & Bresli, 2014) ומעורבות של הלומדים (Bloxham & West, 2004). דווח אף כי סטודנטים חשים יותר נוח מבחינה חברתית להעניק לעמיתיהם משוב איכותני מאשר הערכה כמותית (Cestone, Levine & Lane 2008).

ימה הופך משוב עמיתים אחד לאיכותי יותר מן האחר? הינה שאלה המצריכה חקירה. מחקרים הראו כי משוב הינו יעיל יותר כאשר הוא כולל הצעות לשיפור (Bangert-Drowns et al., 1991; Butler, 1987). משוב המבוסס על הצעות לשיפור מניע לומדים לשנות ולתקן את עבודתם ולפתח הבנה עמוקה יותר של התוכן הלימודי (Coll, Rochema & De Alvarez, Espasa & Guasch, 2012). עם זאת, מחקרים אלו לא הדגישו את הכישורים הקוגניטיביים הנדרשים מצדם של המעריכים – אלו שכותבים את המשובים. סקירת ספרות מעמיקה הראתה כי יש צורך בפיתוח טקסונומיה ברורה המודדת את איכות המשובים תוך שימת דגש על רמות החשיבה הנדרשות לשם ביצועה.

הצורך בביצוע הערכת עמיתים מתעצם בקורסים מקוונים מרובי משתתפים, בהם הציון מבוסס לא רק על מבחנים סגורים, אלא גם על עבודות פתוחות (Barak, Watted & Haick, 2016; Kulkarni et al., 2014). כיום ניתן להבחין בין קורסים מקוונים מסוג MOOC (massive open online course) או SPOC (small open online course) אלפי משתתפים, ללא מחויבות וללא דרישות מוקדמות, או

private online course) המתארים קורסים מקוונים הפונים לקבוצת לומדים פרטית ומצומצמת שנקבעת על פי דרישות מוקדמות. במרבית קורסי ה-MOOC וה-SPOC נעשה שימוש במבחנים סגורים, אלא שאלו אינם מתאימים למשימות לימודיות שמטרתן לפתח חשיבה ברמות גבוהות. על מנת שהלמידה תהיה משמעותית ושהלומד יוכל לפתח חשיבה ביקורתית, יש צורך במשימות פתוחות המעודדות שיתוף ידע (Cevik, 2015; Barak et al., 2016). אלא שבסביבה מקוונת הכוללת מספר סטודנטים רב, הערכת משימות פתוחות על ידי צוות קורס מצומצם הינה משימה בלתי אפשרית (Suen, 2014; Luaces et al., 2015). מכאן, שהפתרון לכך הוא ביצוע הערכת עמיתים מקוונת.

כיום, מרבית קורסי MOOC ו-SPOC כוללים הערכת עמיתים מקוונת בשל העובדה שהיא מפחיתה את עומס העבודה והמשאבים הנדרשים מסגל ההוראה (Suen, 2014; Barak et al., 2016) ואף תורמת לשבירת הבידוד החברתי האופף קורסים הניתנים מרחוק. יחד עם זאת, איכותה עדיין מוטלת בספק. לאור זאת, חוקרים החלו בבחינת הפוטנציאל של שילוב הערכת עמיתים מקוונת בקורסים אלו (Kulkarni et al., 2013; Ashton & Davies, 2015), אך המחקרים בתחום זה עדיין בראשית דרכם.

מטרה ושאלת המחקר

מטרת המחקר הנוכחי הייתה בחינת איכות הערכת העמיתים המקוונת – ציון מספרי ואיכות המשוב – הניתנת בסביבות למידה שונות (כיתת הלימוד, SPOC ו-MOOC). מטרה זו העלתה את שאלת המחקר הבאה:

באיזו מידה סביבת הלמידה, וכפועל יוצא מכך ההיכרות בין הסטודנטים, משפיעה על ציוני ומשובי העמיתים שהם מספקים?

המחקר כלל מדגם של 150 סטודנטים אשר למדו את אותו קורס הנדסי עם אותו צוות הוראה, אך בשלוש סביבות למידה שונות, הנבדלות זו מזו בין היתר מבחינת רמת ההיכרות השוררת בין הסטודנטים. קבוצת הכיתה ($n = 50$) כללה סטודנטים מהטכניון שלמדו באופן פרונטלי. רמת ההיכרות ביניהם הייתה גבוהה מאחר והם נפגשו כל שבוע בכיתה הלימוד. קבוצת ה-SPOC ($n = 50$) כללה סטודנטים מהטכניון אשר לקחו חלק בגרסה המקוונת של הקורס, אך כקבוצה נפרדת. רמת ההיכרות ביניהם הייתה בינונית, שכן הם הגיעו מאותן הפקולטות ולמדו קורסים אחרים יחדיו. קבוצת ה-MOOC ($n = 50$) כללה לומדים ממגוון רחב של מדינות, ללא כל הכרות מוקדמת ביניהם.

קבוצות הכיתה וה-SPOC ייצגו אוכלוסייה די הומוגנית של לומדים מאחר והם מתגוררים באותה המדינה ובעלי רקע לימודי דומה, אך הן נבדלו זו מזו מבחינת סביבת הלמידה שלהן (פנים אל פנים מול מקוונת). מנגד, קבוצת ה-MOOC ייצגה אוכלוסייה הטרוגנית אשר כללה שילוב של לומדים בעלי רקע מגוון מבחינת מגדר (28% נשים, 72% גברים) גיל (טווח של 20-55) ומדינות מוצא. קבוצות ה-MOOC וה-SPOC היו דומות מבחינת סביבת הלמידה (מקוונת), אך נבדלו מבחינת הרקע האישי של הלומדים. רמת ההיכרות אשר שררה בין הסטודנטים בכל אחת מן הקבוצות הייתה פועל יוצא של סביבת הלמידה ממנה הגיעו ושל מאפייניהם הדמוגרפיים והאקדמיים של הלומדים.

כחלק מדרישות הקורס התבקשו הסטודנטים לעבוד בקבוצות של 3-4 סטודנטים על פרויקט הדורש תכנון ננו-חיישן חדשני. לאחר ההגשה הם התבקשו להעריך באופן יחידני את הפרויקטים של לפחות 3 קבוצות עמיתים. ההערכה נעשתה באופן אקראי ובעילום שם וכללה שני שלבים. ראשית, מתן ציון מספרי על פי מחוון ציונים אשר שוקלל לציון מאוני. שנית, מתן הערכה מילולית של משוב מושכל שכולל הצעות לשיפור העבודה. מאחר והסטודנטים מקבוצת הכיתה וה-SPOC היו רשומים לאוניברסיטה באופן רשמי, הם קיבלו ציונים הן מעמיתיהם והן משתי המתרגלות.

המחקר מבוסס על מערך מחקר מסוג ניסויי מבוקר. **ציוני העמיתים** נבחנו על ידי חישוב ממוצע הציונים שהסטודנטים קיבלו מעמיתיהם ומהמתרגלות. ההבדלים בין ממוצעי הציונים שניתנו על ידי הלומדים משלוש קבוצות המחקר נבחנו על ידי ניתוח שונות חד כיווני (ANOVA) והמתאמים בין ציוני הסטודנטים לציוני המתרגלות נמדדו על ידי מבחן פירסון. **משובי העמיתים** נבחנו על ידי חלוקתם למקטעים על פי תוכנם ומתן ניקוד לכל מקטע בהתאם לקטגוריות הקידוד: חיזוק (נק' 1), הצהרה (2 נק'), אימות (3 או 4 נק') והרחבה (5 או 6 נק'). כל קטגוריה מייצגת רמה עולה של איכות משוב ומיומנויות חשיבה נדרשות. שתי הקטגוריות האחרונות אומצו חלקית ממאמרם של Gielen ו-De Wever (2015) ו-Perkins (2003).

חיזוק – קטגוריה המתייחסת להערות כלליות ללא כל תרומה קוגניטיבית לשיפור ולכן קיבלה נקודה 1 בסולם איכות המשוב. הקטגוריה כללה שתי קטגוריות משנה: הערות מסוג חיובי,

המצהירות כי דבר מה נעשה היטב, והערות מסוג שלילי, המצהירות כי דבר מה נעשה בצורה לא מספקת.

הצהרה – הערות הכוללות רשימת תיוג של גורמים הנמצאים או חסרים בעבודה המוערכת ולכן קיבלה 2 נקודות בסולם איכות המשוב. הקטגוריה כללה שתי קטגוריות משנה: הערות מסוג טכני שהתייחסו לעמידה או לאי עמידה בהנחיות הטכניות להגשת המשימה, והערות מסוג סמנטי שכללו שימוש רדוד במושגים מדעיים המוזכרים בעבודה המוערכת.

אימות – הערות הכוללות השוואה בין הדרישות המדעיות של הפרויקט (כפי שהופיעו במחווה) לבין מה שנעשה בפועל. הקטגוריה חולקה לשתי קטגוריות משנה: תקף (3 נק') – עמידה בפרמטרים המדעיים שבמחווה, או לא תקף (4 נק') – אי עמידה בפרמטרים המדעיים שבמחווה.

הרחבה – קטגוריה המתייחסת להערות הכוללות זיהוי של פערים מדעיים בעבודה המוערכת. הקטגוריה חולקה לשתי קטגוריות משנה: הרחבה מסוג מידעי (5 נק') – משובים המספקים מידע נוסף בנוגע להיבטים מדעיים שחסרים בעבודה או הרחבה מסוג הצעתי (6 נק') – משובים המספקים רעיונות והצעות ישימות לשיפור הפרויקט, פתרונות מדעיים אלטרנטיביים, ואו הצעות למחקרי המשך.

שתי הקטגוריות הראשונות בסולם: **חיזוק והצהרה** מצביעות על מאמץ קוגניטיבי מועט מצד המעריך ולפיכך נחשבות למיומנויות חשיבה ברמה נמוכה. קטגוריית **אימות** מצביעה על יכולת אנליטית ולכן על חשיבה ברמה בינונית-גבוהה. קטגוריית **הרחבה** דורשת יכולות של סנתזה והערכה ולכן מצביעה על רמה קוגניטיבית גבוהה. טבלה 1 מציגה דוגמאות לקטעי משוב עבור כל קטגוריה.

טבלה 1. דוגמאות לקטעי משוב על פי ארבע קטגוריות הניתוח (תרגום חופשי)

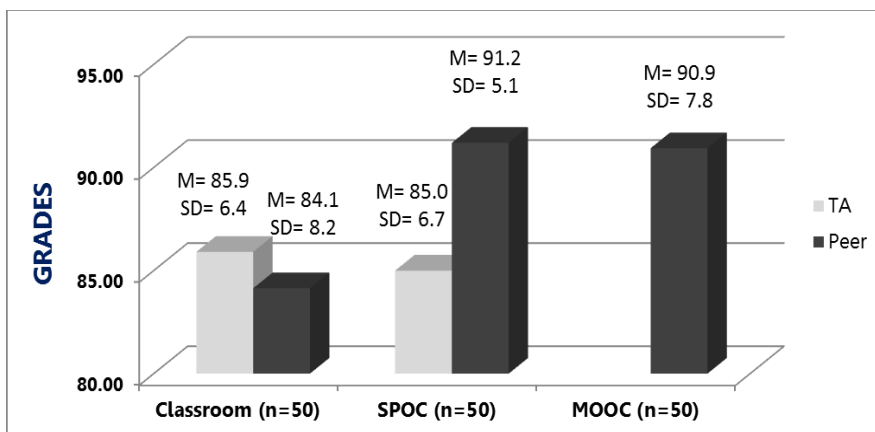
קטגוריה ראשית	קטגוריה משנית	דוגמה למשוב
1. חיזוק	חיובי (1 נק')	• כל הכבוד! (Classroom)
	שלילי (1 נק')	• סך הכל, פרויקט די גרוע (SPOC)
2. הצהרה	טכני (2 נק')	• המושג נחקר ביסודיות עם הפניות טובות (Classroom)
	סמנטי (2 נק')	• שיטת ייצור ננו-החיישן שלך אינה מוזכרת בבירור (MOOC)
3. אימות	תקף (3 נק')	• ננו-צינורית הפחמן המוצגת בפרויקט יכולה לפתור בעיות רפואיות רבות ולכך עשויה להיות השפעה רבה על שלומה של האנושות (MOOC)
	לא תקף (4 נק')	• הרעיון אינו חדשני כבר יותר מ-10 שנים, קבוצת מדענים מאוניברסיטת טוקיו המציאו את רובוט העור האלקטרוני (MOOC)
4. הרחבה	מידעי (5 נק')	• לא הסברת את תהליך ייצור השבב. לדוגמה, עליך להוסיף באיזה סוג של פס ייצור מדובר. נסה לצפות שוב בהרצאה מקוונת מספר 5 (SPOC)
	הצעתי (6 נק')	• היכן אתה מתכוון לשים את החיישן? במעיים של הכלב? שים לב שהאנליט נמצא בדם של הכלב, כך שהייתי מציע לך שהחיישן יקושר למערכת הדם (Classroom)

ממצאים

ציוני עמיתים: השוואה בין שלושת קבוצות המחקר ובדיקת המתאם עם ציוני המתרגלות

מספר העבודות שהסטודנטים העריכו היה: 3 עבודות במוצע בקבוצת הכיתה, 3.60 עבודות במוצע בקבוצת ה-SPOC, ובקבוצת ה-MOOC 5.68 עבודות במוצע. מבחינת הציונים הממוצעים שהסטודנטים נתנו לעמיתיהם, נמצא כי הסטודנטים בקבוצת הכיתה נטו להעניק לעמיתיהם ציונים ממוצעים נמוכים יותר ($M = 84.14$, $SD = 8.25$) בהשוואה לקבוצת ה-SPOC ($M = 91.20$, $SD = 5.06$) או לקבוצת ה-MOOC ($M = 90.94$, $SD = 7.80$). ניתוח שונות חד כיווני הראה שהבדלים אלו מובהקים סטטיסטית ($F(2,147) = 15.48$, $p = .001$, $\eta_p^2 = .174$).

ממוצע ציוני העמיתים של קבוצת הכיתה וה-SPOC נבחנו אל מול ממוצע הציונים שנתנו המתרגלות. הסטודנטים בקבוצת הכיתה העניקו לעמיתיהם ציונים דומים אך נמוכים במקצת מאלו של המתרגלות ($M = 84.14, SD = 8.25$; $M = 85.90, SD = 6.46$). מנגד, הלומדים ב-SPOC העניקו לעמיתיהם ציונים גבוהים מאלו שנתנו המתרגלות ($M = 85.00, SD = 6.77$; $M = 91.20, SD = 5.10$). מבחן פירסון הראה שציוני העמיתים בכיתה נמצאו במתאם בינוני, מובהק סטטיסטית, עם ציוני המתרגלות ($r(50) = .48, p < .001$). לעומתם ציוני קבוצת ה-SPOC נמצאו במתאם נמוך, מובהק סטטיסטית, עם ציוני המתרגלות ($r(50) = .30, p < .005$). הציונים הממוצעים של העמיתים והמתרגלות מוצגים באיור 1.

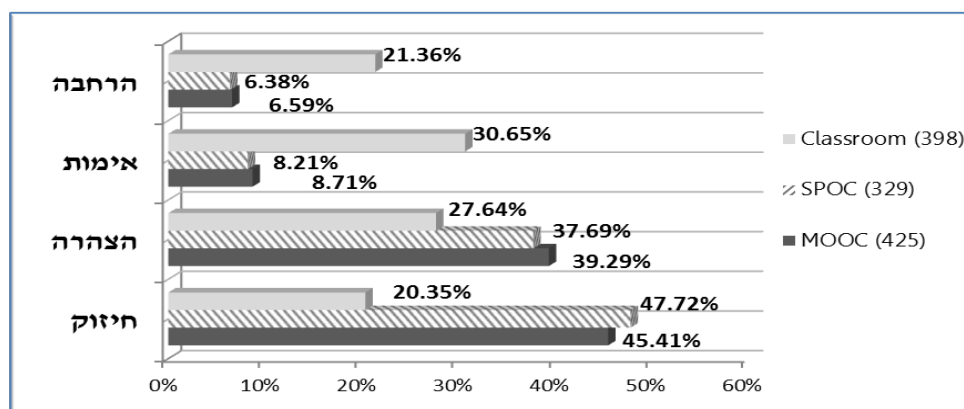


איור 1. הציונים הממוצעים של העמיתים ושל המתרגלות על פי קבוצות המחקר

משובי עמיתים: השוואה בין שלושת קבוצות המחקר

הניתוח שלנו זיהה סך הכול 1152 מקטעי משוב: 398 בקבוצת הכיתה, 329 בקבוצת ה-SPOC ו-425 בקבוצת ה-MOOC. בממוצע, כל סטודנט מקבוצת הכיתה רשם 7.96 מקטעי משוב עבור הערכת עמיתים ($SD = 4.03$). מנגד, כל סטודנט בקבוצת ה-SPOC רשם בממוצע 6.58 מקטעי משוב (ובקבוצת ה-MOOC 8.50 מקטעי משוב ($SD = 6.21$)).

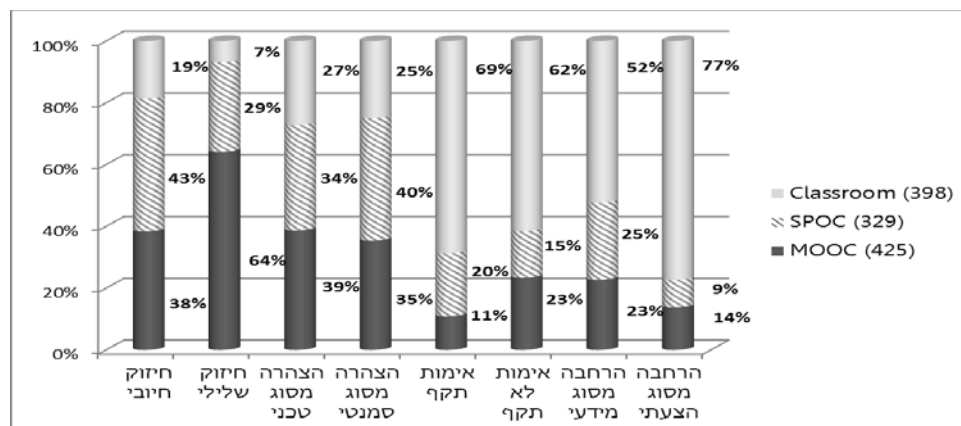
התפלגות מקטעי המשוב בהתאם לארבע קטגוריות הקידוד מוצגת באיור 2. הערות מסוג חיזוק והצהרה, המסווגות כמשוב ברמה נמוכה, היו פחות שכיחות בקרב הסטודנטים בכיתה. הערות מסוג אימות, שסווגו כמשובים באיכות בינונית גבוהה, היו שכיחות יותר בקרב קבוצת הכיתה (30.65%), אל מול סטודנטים בקבוצת ה-SPOC (8.21%) וה-MOOC (8.71%). באופן דומה, הערות מסוג הרחבה, שסווגו כמשובים באיכות הגבוהה ביותר, היו גם הן שכיחות בקרב קבוצת הכיתה (21.36%), בהשוואה לקבוצת ה-SPOC (6.38%) וה-MOOC (6.59%).



איור 2. התפלגות מקטעי המשובים בהתאם לארבע קטגוריות הקידוד

השכיחות היחסית של מקטעי המשובים בהתאם לקטגוריות המשנה מוצגת באיור 3. מבחינת קטגוריות המשנה של חיזוק, נמצא כי הסטודנטים בקבוצת הכיתה העניקו לעמיתיהם אחוז נמוך

מאוד של הערות שליליות (7%). בנוסף, הסטודנטים בקבוצת הכיתה רשמו אחוז גבוה יחסית של משובי אימות מסוג "תקף" (69%) וכן של משובים מסוג הרחבה בקטגוריית משנה "הצעת" (77%).



איור 3. שכיחות יחסית של מקטעי המשובים בהתאם לקטגוריות המשנה

בממוצע, דרוג איכות המשוב שמתתפי קבוצת הכיתה העניקו לעמיתים היה גבוה ($M = 45.23$) בהשוואה לקבוצות ה-SPOC וה-MOOC ($M = 30.51$, $M = 31.54$, בהתאמה). ניתוח שונות חד כיווני לבחינת איכות המשובים הצביע על הבדל מובהק סטטיסטית בין הקבוצות ($F(2,147) = 53.43$, $p < .001$, $\eta_p^2 = .421$).

דין

מחקר זה בחן את הציונים והמשובים שסטודנטים נתנו זה לזה במסגרת פעילות של הערכת עמיתים מקוונת, תוך השוואה בין קבוצת סטודנטים הלומדים בכיתה, בקורס מקוון פרטי (SPOC) ובקורס מקוון רב משתתפים (MOOC). הממצאים מצביעים על העדפה לקבוצת הסטודנטים בכיתה, הן מבחינת ציוני והן מבחינת משובי העמיתים. הסטודנטים מקבוצת הכיתה החמירו עם חבריהם לכיתה והעניקו להם ציונים נמוכים יחסית בהשוואה לציוני העמיתים בקבוצות האחרות ואף בהשוואה לאלו שניתנו על ידי המתרגלות. עם זאת, ציוניהם היו במתאם גבוה יותר עם ציוני המתרגלות. אם כך, סטודנטים אלו הראו יכולת גבוהה יותר להעניק ציונים באופן דומה לזה של "מומחים" בתחום. הממצא עשוי להצביע על כך שסטודנטים שמכירים זה את זה, על אף שההערכה אנונימית, חשים מחויבות גבוהה יותר כלפי עמיתיהם ולכן מתייחסים למשימת הערכת העמיתים ביתר רצינות. ממצא זה עולה בקנה אחד עם הספרות הטוענת כי הערכה מספרית של סטודנטים עשויה להיות לא פחותה מזו הניתנת על ידי סגל ההוראה (Topping, 1998, 2005). מן הצד השני, הסטודנטים בקבוצת ה-SPOC סיפקו לעמיתיהם ציונים ממוצעים גבוהים מאלו שניתנו על ידי המתרגלות. ממצא זה תואם דיווחים של מספר חוקרים כי סטודנטים נוטים להעניק לעמיתיהם ציונים גבוהים מציוני המרצים בשל אמפטיה ורצון לסייע (Sunol et al., 2016). עם זאת, הציונים שהם נתנו לעמיתים היו במתאם נמוך לאלו של המתרגלות.

המחקר הנוכחי עשוי להציע הסבר לפערים בספרות המחקרית בנוגע לציונים הניתנים בהערכת עמיתים. הציונים שנתנו סטודנטים הלומדים בכיתה היו מדויקים יותר מכיוון, שככל הנראה, הם חשים אחריות רבה יותר כלפי עמיתיהם ולכן מתייחסים למשימת ההערכה בצורה רצינית ומשתמשים במחונן בצורה מדויקת. לעומתם, סטודנטים של אותה האוניברסיטה שלמדו מרחוק עשויים לחוש פחות תחושת אחריות כלפי עמיתים ובשל כך לדייק פחות במתן ההערכה.

לעומת זאת, המחקר מצא שהסטודנטים בקבוצת הכיתה נתנו ציונים ומשובים למספר מינימלי של עבודות, בעוד שהסטודנטים ב-SPOC ובייחוד ב-MOOC, ביצעו הערכה למספר רב יותר של עבודות. לתופעה זו הסבר אפשרי הנובע מהעובדה שבשל הריחוק הגאוגרפי האינטראקציה בין הסטודנטים הייתה מוגבלת וכך גם היכולת שלהם לקבל תמיכה מצוות ההוראה. יתכן כי משתתפי קבוצת ה-MOOC הרגישו צורך לספק תמיכה אחד לשני ולתת ציונים ומשוב לכמה שיותר עמיתים, גם אם זה נעשה על חשבון הדיוק בציון ואיכות המשוב.

המחקר הראה כי משובי קבוצות ה-MOOC וה-SPOC היו יותר מפורטים, אך איכותם לא הייתה גבוהה. כ-80% מן ההערות נותרו ברמות חשיבה נמוכות מסוג חיזוק והצהרה. הסטודנטים מקבוצת הכיתה העניקו לעמיתיהם יותר הערות מסוג אימות והרחבה, הנחשבות למשוב ברמה גבוהה, עם

דגש על הצעות קונסטרוקטיביות לשיפור העבודה מבחינה מדעית. הדבר תואם את ממצאיהם של Van der Pol ועמיתיו (2008) שדיווחו כי משוב בו הלומד מציע לעמיתיו רעיונות ממשיים לשיפור העבודה, הינו בעל הפוטנציאל הרב ביותר לקידום למידה משמעותית.

הממצאים הראו שהסטודנטים בקבוצת הכיתה העניקו לעמיתיהם אחוז נמוך של משובים מסוג חיזוק שלילי, בהשוואה לקבוצות האחרות. למרות שהערכת העמיתים נעשתה באופן אנונימי, יתכן כי הסטודנטים מקבוצת הכיתה נמנעו מכתובת תגובות שליליות, בשל תחושת אחריות וחוסר נעימות כלפי עמיתיהם, אותם הם פוגשים על בסיס שבועי. לעומת זאת, יתכן שבשל הריחוק הפיזי ולעיתים בשל הבדלים תרבותיים, לומדי ה-MOOC כתבו הערות חריפות, סרקסטיות ולעיתים פוגעניות.

אחוז גבוה של תגובות מסוג "תקף" בקרב הסטודנטים בכיתה הראה כי הם מקפידים לענות על הדרישות המדעיות של הפרויקט. הדבר עשוי להיות מוסבר על ידי העובדה שכולם היו רשומים לקורס כחלק מתוכנית הלימודים באוניברסיטה והקפידו על דרישות הקורס כדי לסיימו בהצלחה. יתכן גם שההסבר שניתן בכיתה ושאלות ההבהרה ששאלו הסטודנטים בזמן אמת אפשרו הבנה טובה יותר של הדרישות מאשר ההוראות הכתובות באתר הקורסים המקוונים. הדבר ניכר במספר גבוה יחסית של הערות מסוג אימות "לא תקף" שרשמו לומדי ה-MOOC.

על אף שהסטודנטים מקבוצת ה-SPOC היו בעלי מאפיינים דמוגרפיים ואקדמיים דומים לאלו של קבוצת הכיתה ובעלי מטרות לימודיות דומות, איכות המשובים שנתנו הייתה נמוכה. נראה כי קבוצה זו לא ניצלה את הפוטנציאל של למידה היברידית, כלומר, קבלת תמיכה מצוות ההוראה בשעות הקבלה וארגון מפגשים פרונטליים עם חברי קבוצת הלמידה.

מקורות

- Alvarez, I., Espasa, A., & Guasch, T. (2012). The value of feedback in improving collaborative writing assignments in an online learning environment. *Studies in Higher Education*, 37(4), 387-400.
- Ashton, S. & Davies, R. (2015). Using scaffolded rubrics to improve peer assessment in a MOOC writing course. *Distance Education*, 36(3), 312-334.
- Azarnoosh, M. (2013). Peer assessment in an EFL context: attitudes and friendship bias. *Language Testing in Asia*, 3(11). Available at: <http://language-testing-asia.springeropen.com/articles/10.1186/2229-0443-3-11>
- Bangert-Drowns, R. L., Kulik, C. C., Kulik, J. A., & Morgan, M. T. (1991). The instruction effect of feedback in test-like events. *Review of Educational Research*, 6, 218-238.
- Barak, M., & Rafaeli, S. (2004). Online question-posing and peer-assessment as means for Web-based knowledge sharing. *International Journal of Human-Computer Studies*, 61(1), 84-103.
- Barak, M. Watted, A., & Haick, H. (2016). Motivation to learn in massive open online courses: examining aspects of language and social engagement. *Computers & Education*, 94, 49-60.
- Bloxham, S., & West, A. (2004). Understanding the Rules of the Game: Making Peer Assessment as a Medium for Developing Students' Conceptions of Assessment. *Assessment & Evaluation in Higher Education* 29(6), 721-733.
- Butler, R. (1987). Task-involving and ego-involving properties of evaluation: Effects of different feedback conditions on motivational perceptions, interest, and performance. *Journal of Educational Psychology*, 79, 474-482.
- Cestone, C. M., Levine, R. E. & Lane, D. R. (2008). Peer Assessment and Evaluation in Team-based Learning. *New Directions for Teaching and Learning*, 116, 69-78.
- Cevik, Y. D. (2015). Assessor or assessee? Investigating the differential effects of online peer assessment roles in the development of students' problem-solving skills. *Computers in Human Behavior*, 52, 250-265.
- Coll, C., Rochera, M. J., & De Gispert, I. (2014). Supporting online collaborative learning in small groups: Teacher feedback on learning content, academic task and social participation. *Computers & Education*, 75, 53-64.

- Falchikov, N. (1995). Peer feedback marking: Developing peer assessment. *Innovations in Education & Training International*, 32(2), 175-187.
- Falchikov, N., & Goldfinch, J. (2000). Student peer assessment in higher education: A meta-analysis comparing peer and teacher marks. *Review of Educational Research*, 70(3), 287-322.
- Gielen, M., & De Wever, B. D. (2015). Structuring peer assessment: Comparing the impact of the degree of structure on peer feedback content. *Computer in Human Behavior*, 52, 315-325.
- Hanrahan, J., & Isaacs, G. (2001). Assessing self- and peer assessment: The students' views. *Higher Education Research and Development*, 20(1), 53-70.
- Harris, J. (2011). Peer Assessment in Large Undergraduate Classes: An Evaluation of a Procedure for Marking Laboratory Reports and a Review of Related Practices. *Advances in Physiology Education* 35(2), 178-187.
- Kearney, S., Perkins, T. & Kennedy-Clark, S. (2016). Using self- and peer-assessments for summative purposes: analysing the relative validity of the AASL (Authentic Assessment for Sustainable Learning) model. *Assessment & Evaluation in Higher Education*, 41(6), 840-853.
- Kulkarni, C., Wei, K. P., Le H., Chia, D., Papadopoulos, K., Cheng, J., Koller, D., & Klemmer, S. (2013). Peer and Self Assessment in Massive Online Classes. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 20(6), Article 33, pp. 1-31.
- Kulkarni, C., Socher, R., Bernstein, M., & Klemmer, S. (2014). *Scaling short-answer grading by combining peer assessment with algorithmic scoring*. Proceedings of ACM Learning at Scale. March, 2014, pp. 99-108.
- Love, K. G. (1981). Comparison of peer assessment methods: reliability, validity, Friendship bias, and user reaction. *Journal of Applied Psychology*, 66, 451-457.
- Li, H., Xiong, Y., Zang, X., Kornhaber, M. L., Lyu, Y., Chung K. S. & Suen, H. K. (2016) Peer assessment in the digital age: a meta-analysis comparing peer and teacher ratings, *Assessment & Evaluation in Higher Education*, 41(2), 245-264. DOI: 10.1080/02602938.2014.999746
- Li, L. & Gao, F. (2016) The effect of peer assessment on project performance of students at different learning levels. *Assessment & Evaluation in Higher Education*, 41(6), 885-900. DOI: 10.1080/02602938.2015.1048185
- Liu, N. F. & Carless, D. (2006). Peer Feedback: The Learning Element of Peer Assessment. *Teaching in Higher Education*, 11(3), 279-290.
- Luaces, O., Di'ez, J., Alonso-Betanzos, A., Trocoso, A., & Bahamonde, A. (2015). A factorization approach to evaluate open-response assignments in MOOCs using preference learning on peer assessments. *Knowledge-Based Systems*, 85, 322-328.
- McConlogue, T. (2015). Making judgements: investigating the process of composing and receiving peer feedback. *Studies in Higher Education*, 40(9), 1495-1506.
- Morahan-Martin, J. (1996). Should peers' evaluations be used in class projects? Questions regarding reliability, leniency, and acceptance. *Psychological Report*, 78,1243-1250.
- Nicol, D., Thomson, A. & Breslin, C. (2014). Rethinking Feedback Practices in Higher Education: A Peer Review Perspective. *Assessment & Evaluation in Higher Education*, 39(1), 102-122.
- Perkins, D. (2003). *King Arthur's Round Table: How Collaborative Conversations Create Smart Organizations*. Hoboken, NJ: John Wiley & Sons.
- Price, M., Rust, C., O'Donovan, B., Handley, K. & Bryant, R. (2013). *Assessment literacy: The foundation for improving student learning*. Oxford: The Oxford Centre for Staff and Learning Development.
- Sadler, P. M., & Good, E. (2006). The impact of self-and-peer-grading on student learning. *Educational*, 11(1), 1-31.

- Sitthiworachart, J., & Joy, M. (2008). Computer support of effective peer assessment in an undergraduate programming class. *Journal of Computer Assisted Learning*, 24(3), 217-231. doi:10.1111/j.1365-2729.2007.00255.
- Snowball, J. D., & Mostert, M. (2013). Dancing with the Devil: Formative Peer Assessment and Academic Performance. *Higher Education Research & Development*, 32 (4), 646-659.
- Suen, H. K. (2014). Peer Assessment for Massive Open Online Courses (MOOCs). *The International Review of Research in Open and Distance Learning*, 15(3), 312-327.
- Suñol, J. J., Arbat, G. Pujol, J., Feliu, L., Fraguell, R. M. & Planas-Lladó. A. (2016). Peer and self-assessment applied to oral presentations from a multidisciplinary perspective. *Assessment & Evaluation in Higher Education*, 41(4), 622-637.
- Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*, 68(3), 249-276.
- Topping, K. (2005). Trends in Peer Learning. *Educational Psychology*, 25(6), 631-645.
- Van der Pol, J., Van den Berg, B.A.M., Admiraal, W.F. & Simons, P.R.J. (2008). The nature, reception, and use of online peer feedback in higher education. *Computers & Education*, 51, 1804-1817.