

ג כסלו תשפ"א 19.11.2020

## דיוק, אמינות ומהימנות – מה בעצם ההבדל?

מאת: ד"ר עומר חורש

**יעוץ ועריכה: אורה הירש, שרה ורטהימר, אפרת לינק, ד"ר פרחי וקסמן, תמי בן-דור**

דיוק, אמינות ומהימנות הם מושגים שתלמידים מרבים להשתמש בהם באופן שגוי במהלך תכנון ניסוי או בהתייחסות למרכיבי חקר, כגון בהסבר החשיבות של שמירה על גורם קבוע במערכת הניסוי או בהסבר החשיבות שיש בביצוע חזרות או ריבוי פריטים בניסוי. דוגמאות לטעויות נפוצות בתשובות תלמידים:

"חשוב לבצע חזרות או ריבוי פריטים על מנת לוודא שהתוצאות מדויקות".

"חשוב לשמור על גורמים קבועים כדי לוודא שהתוצאות אמינות".

"חשוב לבצע חזרות כדי לוודא שהתוצאות אמיתיות".

במאמר זה נכיר את המושגים אמינות, דיוק, מהימנות ונבהיר את ההבדל ביניהם, כל זאת במטרה לשרש תפיסות שגויות הנפוצות בקרב תלמידים. בנוסף, נתייחס גם למושג תוקף המסקנות.

נתחיל דווקא ב**אמינות**, המושג הכי פחות רלוונטי מבין השלושה בהסבר החשיבות של מרכיבים בניסויים בביולוגיה הנערכים בבתי הספר.

אמינות על פי מילון אבן שושן היא "אותנטיות, אמיתיות, נאמנות למקור, שיש בו יסוד לכך שלא יכזיב", הגדרה הנבחנת לרוב באמירת אמת או שקר. למעשה, במובן רחב יותר, אמינות היא המידה בה אפשר להאמין לדבר מה. דיווח אמין הוא דיווח שהוקפד שלא יהיו בו שקרים או שגיאות, וכן שלא יושמטו מהדיווח תכנים באופן סלקטיבי ומניפולטיבי רק כדי ליצור הטיה בקרב ציבור הנחשף אליו.

בניסויים שתלמידים מבצעים הם מדווחים על התוצאות שקיבלו, לא מוסיפים או משמיטים תוצאות (עלינו להקפיד על כך ולחנך אותם לדרך זו). אנו לא מסתירים תוצאות חריגות אלא מנחים את התלמידים להתייחס אליהן, ולמצוא להן הסברים. אם הם ינהגו כך התוצאות שיקבלו בניסוי תהיינה אמיתיות ולא שקריות, גם אם הם לא ביצעו חזרות, או לא הקפידו על שמירת גורמים קבועים. לכן אין כל משמעות לכך שהתלמידים יוסיפו ויטענו כי התוצאות אמינות בזכות חזרות על הניסוי או ריבוי פריטים.

בנוסף, בתחומי המדע משתמשים במושג **אמינות (Reliability)** כדי לקבוע את ההסתברות שמכשיר או מערכת יתפקדו באופן תקין. במקרה זה המושג רלוונטי למכשירי המדידה בהם אנו משתמשים, למשל פיפטור, ספקטרופוטומטר, מד pH, מד מוליכות. התלמידים לא נדרשים לבדוק זאת במהלך הניסוי, והם יוצאים מנקודת הנחה כי השתמשו במכשיר תקין ומכיל, כלומר אמין. המקרה היחיד שאפשר להשתמש במושג אמינות הוא כאשר יש צורך להסביר תוצאות חריגות או למצוא הסברים

אחרים לתוצאות הניסוי שלא תאמו את ההשערה. ניתן לשער כי החריגות נבעו מהיעדר אמינות של מכשיר המדידה.

ובכן, נותרנו עם המושגים "דיוק" ו"מהימנות". נבחן כעת את ההבדל ביניהם :

**דיוק (Accuracy)** הוא מידת הקרבה של הערך הנמדד לערך האמיתי. הערך האמיתי נקבע מראש על ידי החוקר, או בניסויים קודמים (לרוב מתפרסם בספרי נתונים, בפרוטוקול לביצוע ניסוי או בספרות המדעית).

דיוק קשור בראש ובראשונה למידת הרגישות של מכשיר המדידה. ככל שהמכשיר רגיש יותר, כלומר מסוגל למדוד בערכים קטנים יותר, כך הסיכוי שתתקבל תוצאה קרובה לערך האמיתי היא גבוהה יותר. למשל, אם עלינו למדוד מסה של 2.015 גרם ונשתמש במד מסה המודד ביחידות של גרם בלבד, סביר כי נקבל תוצאה לא מדויקת. לעומת זאת אם נשתמש במד מסה המודד ביחידות מיליגרם ומציג תוצאות בערכים של עד שלוש ספרות לאחר הנקודה, נוכל לקבל תוצאה מדויקת יותר. דיוק במדידות, כגון של מסה ונפח בהכנת תמיסות, רלוונטי בעבודת הליבונט או בעבודת התלמיד במסגרת הביוחקר. דוגמה נוספת בה ניתן להתייחס לדיוק היא הקפדה על מדידה של אורך שורשון ונצרון או של מרחק, למשל של נוזל המתקדם בפיפטה.

היעדר דיוק נובע לא רק מרגישות מכשיר המדידה אלא גם משגיאות שיטתיות, כגון בעיה במכשיר המדידה או בעין החוקר הרושם את הנתונים. למעשה, תוצאות לא מדויקות נובעות משגיאות שיטתיות המרחיקות את הערך המתקבל מן הערך האמיתי. הדיוק קשור באופן מסוים לאמינות התוצאות, שכן אם מכשיר המדידה אינו תקין או אינו מכויל, כלומר לא אמין, סביר כי יתקבלו תוצאות לא מדויקות עם סטייה דומה בכל המדידות. אולם, האם ריבוי פריטים או חזרות על הניסוי או שמירה על גורמים קבועים הם אלה שישפיעו על מידת הדיוק של תוצאות או על אמינות מכשיר המדידה, בהנחה שאנחנו משתמשים באותה שיטת מדידה ובאותו מכשיר?

התשובה היא לא, וכאן טמונה התפיסה השגויה בקרב תלמידים.

**חזרות על הניסוי** מאפשרות לוודא כי התוצאות חוזרות על עצמן ואינן מקריות. אין בהכרח בחזרות על הניסוי כדי להעיד על רמת הדיוק הנובעת מרגישות במכשיר או על מידת האמינות הנובעת מתקנותיו. בנוסף, המושגים אמינות ודיוק אינם רלוונטיים להסבר החשיבות בשמירה על גורמים קבועים בניסוי, למרות שבשמירה על גורמים קבועים רצוי להיות מדויקים ודיווח על גורמים קבועים מעיד על אמינות. שמירה על גורמים קבועים מטרתה אחת, לוודא **שרק** המשתנה הבלתי תלוי משפיע על המשתנה התלוי, בהנחה שגם גורמים אלה משפיעים על המשתנה התלוי.

המושג **מהימנות (Reproducibility)** מתייחס לשגיאות אקראיות בניסוי, ומתייחס למידת הקרבה של התוצאות החוזרות. ראוי לציין שככל שסטיית התקן קטנה יותר מהימנות התוצאות גבוהה יותר. למעשה, תוצאות יכולות להיות שגויות או לא מדויקות אך מהימנות, כלומר חוזרות על עצמן. אם כן, מטרת החזרות על הניסוי או ריבוי הפריטים בניסוי היא להגדיל את מהימנות התוצאות, כלומר לוודא שאינן מקריות.

נעזר בטבלה 1 (עמ' 4) בה מוצגות תוצאות טיטרציה של תמיסה בה מומס  $\text{CO}_2$ , עם תמיסת בסיס הנתרן בנוכחות אינדיקטור. נתון שריכוז ה- $\text{CO}_2$  בתמיסה הוא  $0.0005\text{M}$  ודרושים, על פי ידע קודם,  $2.5$  מ"ל בסיס הנתרן על מנת לטטר כמות זו של  $\text{CO}_2$  בנפח התמיסה הנתון. המדידה נערכה בנוכחות אינדיקטור מאותו הסוג אך ממקורות שונים (מספקים שונים).

על החוקר להגדיר מראש עד לאיזה מרחק של הממוצע מהתוצאה האמיתית התוצאות עדיין ייחשבו מדויקות במידת סבירות גבוהה, וכן להגדיר את טווח פיזור הערכים סביב הממוצע, כך שהתוצאות ייחשבו מהימנות. קביעות אלו נערכות, לרוב, באמצעות מבחנים סטטיסטיים והגדרות נוספות אשר אינם נדרשים מהתלמיד בבחינת הברורות המעשיות, אולם בעבודות הביוחקר בהחלט ניתן לקבוע ערכים אלה, ולהתייחס ביתר כובד ראש לנושא הדיוק והמהימנות בדיון של העבודה. בניסוי שתוצאותיו מוצגות בטבלה 1 נערכו חמש חזרות עבור כל אחד מהאינדיקטורים. נקבע (הגדרת החוקר לצורך העניין) כי מרחק של הממוצע ב- $0.3$  מ"ל או יותר מהערך האמיתי ( $2.5$  מ"ל) מעידה על תוצאות לא מדויקות, וכן, כי סטיית תקן גבוהה מ- $0.3$  מ"ל (מעל  $12\%$  יחסית לערך המדויק/הממוצע) מעידה כי התוצאות מפורזות סביב הממוצע באופן המעיד על כך שאינן חוזרות על עצמן, כלומר אינן מהימנות.

**טבלה 1: מדידה של נפח בסיס הנתרן שיש להוסיף לתמיסה, עד לשינוי בצבע אינדיקטור ממקורות שונים.**

| ממוצע נפח<br>בסיס הנתרן<br>(מ"ל) | נפח בסיס<br>הנתרן (מ"ל)<br>בכל חזרה | מקור<br>האינדיקטור |
|----------------------------------|-------------------------------------|--------------------|
| 2.4                              | 2.5                                 | א <sup>1</sup>     |
|                                  | 2.5                                 |                    |
|                                  | 2.5                                 |                    |
|                                  | 2.0                                 |                    |
|                                  | 2.5                                 |                    |
| 7.3                              | 5.5                                 | ב                  |
|                                  | 9.2                                 |                    |
|                                  | 4.3                                 |                    |
|                                  | 6.7                                 |                    |
|                                  | 10.9                                |                    |
| 9.5                              | 9.5                                 | ג                  |
|                                  | 9.4                                 |                    |
|                                  | 9.5                                 |                    |
|                                  | 9.6                                 |                    |
|                                  | 9.4                                 |                    |
| 2.6                              | 2.5                                 | ד                  |
|                                  | 3.1                                 |                    |
|                                  | 2.4                                 |                    |
|                                  | 2.0                                 |                    |
|                                  | 3.2                                 |                    |

<sup>1</sup> שימו לב לכך שהתוצאה 2.0 מ"ל אמנם נחשבת ללא מדויקת וחריגה, אולם אנו מתייחסים לממוצע נפח בסיס הנתרן מבלי להעלים את התוצאה החריגה כלומר מקפידים על דיווח אמין של תוצאות. דוגמה זו ממחישה את היתרון בהצגת ממוצע, המקטין את ההשפעה של תוצאות חריגות. דוגמה נוספת קיימת בתוצאות ממקור ד.

ניתן להסיק, על פי הדוגמה שבטבלה, שחזרות על הניסוי אינן גורמות לכך שהתוצאות תהיינה יותר מדויקות. רק במקרים בודדים, חזרות על הניסוי גורמות לכך שהתוצאות תהיינה יותר מדויקות. במקרים אלו, ניתן לטעון שהחוקר משתפר מחזרה לחזרה ביכולת המדידה שלו, ולכן הוא מדייק יותר בקביעת הערך, אך לרוב מכשיר המדידה הוא הקובע את הערך. אם בידי החוקר יש קריטריונים לדיוק שהוגדרו מראש, הוא יכול לקבוע אם התוצאות מדויקות כפי שהוצג בטבלה. אולם בפועל, בבחינת הבגרות המעשית, חסר מידע אמין לגבי הערך האמיתי בניסוי (נקודת הייחוס) לכן אין משמעות לקביעה כי התוצאות מדויקות.

לעתים, במסגרת הביוחקר משתמשים בשיטת מדידה המוכרת בספרות המדעית. במקרים אלו<sup>2</sup> ניתן להשוות בין הערכים שהתקבלו בניסוי ובין הערכים שבספרות. הערך מהספרות יקבע כערך האמיתי ונוכל להסיק על פיו מהי מידת הדיוק של התוצאות בניסוי שערכו התלמידים.

חוקר יכול לקבוע האם תוצאות מהימנות על פי קריטריונים סטטיסטיים. בבחינת בגרות מעשית התלמידים אינם משתמשים בקריטריונים סטטיסטיים לכן הם אינם יכולים להסיק מהתוצאות שיקבלו לגבי מידת המהימנות של הניסוי, כלומר עד כמה התוצאות חוזרות על עצמן. למרות זאת, צפוי שהתלמיד יבין שמטרת החזרות היא לוודא שתוצאות הניסוי מהימנות, כלומר אינן מקריות<sup>3</sup>.

**תוקף מסקנות** הוא מושג נוסף שבו נעשה לעתים שימוש שגוי בתשובות תלמידים בבחינת הבגרות או בדיון במסגרת הביוחקר. דוגמה לטעות נפוצה בתשובות תלמידים:

"חשוב לבצע חזרות על הניסוי על מנת לוודא שהמסקנות תקפות".

מידת התקפות של המסקנות תלויה במערך תקין של הניסוי. ניתן לטעון כי מסקנות מהניסוי הן תקפות בתנאי שבניסוי נערכו כל הבקורות המתאימות לשלילת הסברים חלופיים. כמו כן, ניתן לקבוע שהמסקנות תקפות לתנאי הניסוי (כגון האורגניזם וטווח הערכים שנבדק) ולא ניתן להכליל אותן באופן גורף. אם נקפיד על כל אלה, הם יעידו על תקפות המסקנות, ואילו חזרות לא יעידו על תקפות אלא על מהימנות.

לסיכום, מבין שלושת המושגים המשמשים לביסוס מסקנות של ניסוי - אמינות, דיוק ומהימנות, ברוב המקרים, כולל בבחינת הבגרות, רק המושג מהימנות הוא ענייני למה שביכולתו של התלמיד להסביר, ותקפות עניינה אחר.

מידע נוסף, מאמרים, דגמי הוראה וטעויות המשגה ותרגול בנושא דיוק, מהימנות<sup>4</sup>, אמינות ותוקף ניתן למצוא באתר המרכז הארצי למורי הביולוגיה ומדעי הסביבה [בקישור זה](#)<sup>5</sup>, או בסריקת הקוד:

<sup>2</sup> רק כאשר תנאי הניסוי והאורגניזם הנבדק הם זהים.

<sup>3</sup> התוספת המודגשת חשובה משום שהיא מעידה על כך שהתלמיד מבין את משמעות המושג מהימנות. על פי תכנית הלימודים התלמיד אינו חייב להשתמש במושג מהימנות.

<sup>4</sup> נהוג להשתמש במושג אמינות ובמושג מהימנות בהקשר למקורות מידע. ראו [הנחיות לביוחקר תשע"ז](#) עמ' 30. במקרה זה יש שימוש באותו מושג, [בהקשר אחר ובמשמעות שונה](#). כדאי להסב את תשומת לב התלמידים להבדלים אלה.



---

<sup>5</sup> שימו לב לכך שהגדרת המושג מהימנות בדגם ההוראה "הבנת חשיבות החזרות וריבוי פריטים למהימנות התוצאות" היא רחבה וכוללת בתוכה גם את המושג "דיוק", אך אין בהם סתירה לדברים הנכתבים במאמר זה.