



הנחיות לסקרים: נספח א' – הרחבת מושגים

אפריל, 2025

ניסן, תשפ"ה

הנחיות לסקרים: נספח א' – הרחבת מושגים

תוכן עניינים

פתיח	- 4 -
תכנון הדגימה	- 4 -
אמידה באמצעות מדגמים הסתברותיים	- 5 -
אוכלוסייה	- 6 -
אוכלוסיית המטרה	- 6 -
אוכלוסיית הסקר	- 6 -
מסגרת הדגימה	- 7 -
ייצוג	- 7 -
הטיית כיסוי	- 7 -
הטיות בדגימה	- 8 -
שיטות לדגימה הסתברותית	- 8 -
דגימה אקראית פשוטה (Simple Random Sample)	- 8 -
דגימה שיטתית (Systematic Sampling)	- 8 -
דגימת שכבות (stratified sampling)	- 8 -
דגימת אשכולות (cluster sampling)	- 9 -
אמידה באמצעות מדגמים לא-הסתברותיים	- 10 -
רקע	- 10 -
שיטות לדגימה לא-הסתברותית	- 11 -
דגימה מותאמת (sample matching)	- 12 -
דגימה באמצעות רשת קשרים (network sampling)	- 12 -
דגימת נוחות (convenience sampling)	- 14 -
הטיות ייצוג בשיטות לדגימה לא-הסתברותית	- 15 -
שיטות להתאמה ותיקון בשלב הדגימה	- 15 -

- 17	דגימה מתוך פנלים
- 20	תכנון כלי מדידה ושיטות איסוף
- 20	תוקף ומהימנות של המדידה
- 21	תוקף תוכן (content validity)
- 21	תוקף מבנה (construct validity)
- 22	תוקף תלוי קריטריון (criterion related validity)
- 22	מהימנות פנימית (internal reliability)
- 23	מהימנות מבחן חוזר (repeatability or test-retest reliability)
- 23	מקורות להטיה בפיתוח שאלון
- 24	הטיות שמקורן בשאלון ובתהליך המענה
- 26	שימוש במראיין חליפי
- 28	מקורות להטיה באיסוף הנתונים
- 28	שיטת איסוף הנתונים
- 30	שימוש במראיינים
- 31	מתן תגמול לעידוד ההשבה
- 32	התמודדות עם הטיות במדידה ובאיסוף
- 32	בדיקות מקדימות
- 34	בדיקות בשדה
- 35	בדיקות לאחר עריכת הסקר
- 36	המלצות לעיצוב השאלון ולניסוח שאלות הסקר
- 40	תכנון עיבוד וניתוח
- 40	הכנת הנתונים לעיבוד
- 40	חוסר השתתפות בסקר
- 41	אמות מידה להשתתפות בסקר
- 42	איתור מקורות להטיית אי-השבה
- 43	מדדי תפוקה (outcome rates) למדגמים הסתברותיים

- 44 -	 שיעור השבה (response rate)
- 45 -	 שיעור שיתוף פעולה (cooperation rate)
- 46 -	 שיעור סרבנות (refusal rate)
- 46 -	 שיעור יצירת קשר (contact rate)
- 47 -	 מדדי תפוקה בפנלים הסתברותיים ולא-הסתברותיים
- 47 -	 שיעור גיוס (recruitment rate)
- 48 -	 שיעור השלמת פרופיל (profile rate)
- 48 -	 שיעור השלמת שאלון (completion rate)
- 49 -	 שיעור השבה כולל (cumulative response rate)
- 49 -	 שיטות להתאמה ותיקון במדגמים לא-הסתברותיים
- 50 -	 שיטות לכיול גלובלי (global adjustment)
- 51 -	 שיטות לכיול מותאם-תוצאה (outcome-specific adjustment)
- 52 -	 תוקף הממצאים
- 52 -	 תוקף חיצוני (external validity)
- 53 -	 תוקף פנימי (internal validity)
- 55 -	 אתיקה
- 55 -	 פרטיות וסודיות
- 55 -	 סיכון לפגיעה בפרטיות
- 56 -	 הסכמה מדעת
- 57 -	 מכשולים והטיות שקשורים בהסכמה מדעת
- 58 -	 מאמצים לצמצום הפגיעה בפרטיות ולשמירה על סודיות הנתונים
- 60 -	 נטל השבה
- 61 -	 יסודות להנחיות אתיקה
- 61 -	 גופים סטטיסטיים
- 67 -	 גופי מחקר ואקדמיה
- 71 -	 מונחון

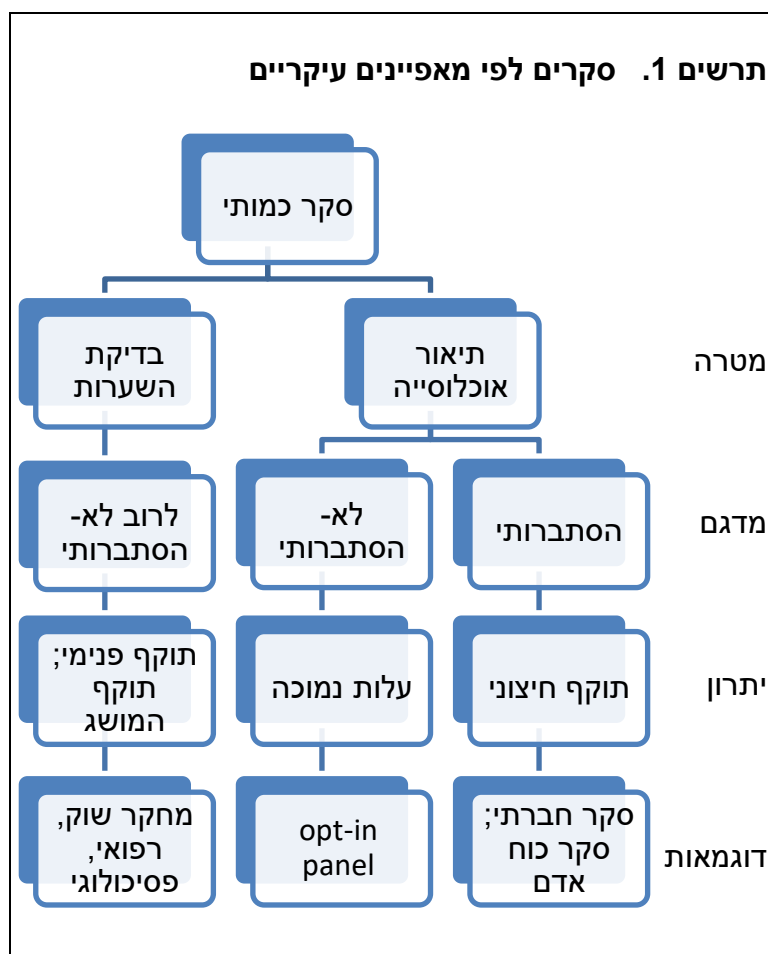
נספח מושגים

פתיח

מטרתו של הנספח היא לתת הרחבה על מושגים ונושאים בסיסיים שמופיעים במסמך ההנחיות. הרחבה זו נועדה לתת תמונה על מגוון שיטות מרכזיות, **כאשר על מזמין הסקר ו/או מבצעו מוטל להחליט אלו מבין השיטות ראויים להיבחר לסקר**, בהתאם לעניין ולצורך. יש להדגיש שרק מיעוט מכל מה שרשום כאן יבוא לידי ביטוי בסקר נתון.

בישראל מבוצעים סקרים רבים על ידי גופים ממשלתיים, האקדמיה, מכוני מחקר ומכונים פרטיים. הסקרים מכסים תחומים רבים – כלכלה, חברה, חינוך, רווחה, ועוד, ונעשים לצרכים מגוונים – חזוי דמוגרפי, תכנון תכניות התערבות, ניתוח צרכים של צרכנים, ועוד. לרוב, איסוף המידע נעשה באמצעות מדגם. המידע יכול להיאסף באופן חד-פעמי או במספר פעמים, מאותו מדגם של משיבים או ממדגמים שונים, ובאופנים שונים (אינטרנט, טלפון, ריאיון פנים אל פנים).

תכנון הדגימה



ניתן להבחין בין שני סוגי סקרים: סקרים לשם תיאור של אוכלוסייה וסקרים שמטרתם לבחון השערות כיצד גורמים שונים קשורים לתוצאות שונות (ראו תרשים 1). סקרים שמתארים אוכלוסייה מספקים אומדנים לפרמטרים (למשל סך הכול, ממוצע, חציון, אחוזים) על מגוון משתני תוצאה (למשל, שכר, גיל, תמיכה במועמד פוליטי וכדומה). ההיבט המרכזי בסקרים אלה הוא התוקף החיצוני של הסקר, כלומר, מדת הייצוג (representation) שלו את האוכלוסייה. קובעי מדיניות במשרדי ממשלה הם דוגמה לצרכנים טיפוסיים של סטטיסטיקה

זו לעומת זאת, מטרתם של סקרים שבודקים השערות אינה בהכרח לייצג אוכלוסייה, אלא לחדד

ולזכר מושגים ולבסס מודלים סיבתיים. רבים מסקרים אלה נעשים על ידי האקדמיה ומכוני מחקר בתחומים פסיכולוגיים, חינוכיים, רפואיים וכדומה, ואף במחקרי שוק.

על מנת שסקר יאפשר ייצוג של אוכלוסיית המטרה, נדרש שהמדגם יעמוד בתנאים האלה:

1. היחידות שנדגמו (אנשים, משקי בית וכדומה) הן בנות החלפה עם יחידות שלא נדגמו, כלומר, הן חולקות את אותם המאפיינים (כל המשתנים שנמדדים וכל משתנה מתערב אפשרי שעשוי להיות קשור אליהם);
2. לא יהיו חלקים מאוכלוסיית המטרה בעלי מאפיינים ייחודיים (אשר הם רלוונטיים לנושא הסקר) שיושמו באופן שיטתי מהמדגם;
3. הרכב המדגם, מבחינת המאפיינים שניתנים למדידה ואשר רלוונטיים ליעדי הסקר, יהיה תואם להרכב אוכלוסיית המטרה, או שניתן יהיה להתאים אותו כך שיהיה תואם לאוכלוסיית המטרה מבחינת הרכבו.

סקרים המבוססים על דגימה הסתברותית (probabilistic sampling) ואלה המבוססים על דגימה לא-הסתברותית (non-probabilistic sampling) נבדלים אלה מאלה בסוג ההצדקה שניתנת למידת הדיוק (לחילופין, במידת הסיכון לטעות). מדגם הסתברותי מתבסס על תיאוריית הדגימה ההסתברותית, אשר נסמכת על עקרונות מתמטיים ידועים. דגימה לא-הסתברותית, לעומת זאת, מצדיקה את מידת הדיוק במודלים; אלה כוללים הנחות שונות לגבי שימוש במשתנים שנועדו לשלוט בהטיות ולהתאים את המדגם להרכב האוכלוסייה.

חלקו הראשון של המסמך יתאר עקרונות בסיסיים של דגימה הסתברותית. חלקו השני של המסמך יתייחס למצבים שבהם יש הכרח להשתמש בדגימה לא-הסתברותית, ויכלול סוגים שונים לתיקון הטיות בדגימה. חלקו השלישי של המסמך יתאר עריכת סקרים באמצעות פנלים.

בטרם נפנה לתאר את סוגי המדגמים השונים, נבקש להפנות את תשומת לב הקורא לקיומם של מונחים שונים, שיש לשים לב להקשר שבו הם מוזכרים. למשל, "מדגם מייצג" בהקשר של דגימה הסתברותית שונה במשמעותו בהקשר של דגימה לא-הסתברותית: הראשון, נוצר מעצם פעולת הדגימה ואילו השני נבנה באופן כזה שייצג (ישקף) התפלגויות נתונות של משתנים רלוונטיים באוכלוסייה. דוגמה אחרת היא "מדגמי פנל", שעשויים להיות הסתברותיים או לא-הסתברותיים, ואיסוף הנתונים עשוי להיות באמצעות הטלפון או האינטרנט. דוגמה נוספת היא "מדגם מכסה", שעשוי להיחשב כמדגם נוחות או כמדגם מתואם, בהתאם לסוג המשתנים, למספרם ולמידה שבה הם משקפים את התפלגות המשתנים באוכלוסייה. כל אלה יפורטו בהמשך.

אמידה באמצעות מדגמים הסתברותיים

סקרים שמתארים את האוכלוסייה, כוללים משתני תוצאה מגוונים. למשל, סקרי כוח אדם מטעם הלשכה המרכזית לסטטיסטיקה מתארים היבטים כלכליים-חברתיים של אוכלוסיית מדינת ישראל; סקרים פוליטיים וסקרי בחירות נותנים תחזיות לגבי שיעור המצביעים, מקרב אוכלוסיית המדינה, למפלגה כזו או אחרת; סקרי שוק בודקים את שביעות הרצון של לקוחות חברה ממוצר או משירות. הסקרים מוסרים אומדנים על מספר, שיעור, ממוצע, התפלגות, או כל פרמטר אחר בהתייחס

לאוכלוסייה שהוגדרה על ידי מזמין הסקר. במקרים כאלה האומדנים שמתקבלים מהסקר על האוכלוסייה צריכים להיות מדויקים ככל האפשר, כאשר הדגימה המתאימה ביותר לכך היא באמצעות דגימה הסתברותית. דגימה הסתברותית היא שם כולל לפרוצדורות לבחירת מדגם, שבהן היחידות שבמדגם נבחרות בצורה אקראית (רנדומלית) ממסגרת דגימה, ולכל יחידה יש סיכוי חיובי וידוע להיכלל במדגם. לעומת זאת, כאשר בחירת המדגם אינה נעשית בצורה אקראית, או שהסיכוי לבחירה אינו ידוע, מדובר בדגימה לא-הסתברותית (ראו בהמשך).

אוכלוסייה¹

אוכלוסיית המטרה

אוכלוסיית המטרה היא קבוצת יחידות שעליהן החוקר מעוניין להסיק מסקנות באמצעות הסטטיסטיקה שנאספת מהמדגם. אוכלוסיית המטרה היא לרוב סופית בגודלה (לפחות באופן תיאורטי ניתן למנות אותה), היא תחומה בזמן ובמקום וניתנת למדידה; כל אלה היבטים שמאפשרים להבין את משמעות הסטטיסטיקה שנאספה בסקר ולערך שיחזור של הסקר.

אוכלוסיית הסקר

אם איסוף המידע נעשה באמצעות סקר, האוכלוסייה בפועל שממנה ייאסף המידע בסקר נקראת אוכלוסיית הסקר (למשל, אוכלוסיית המטרה יכולה להיות הכלבים הנמצאים בבעלות פרטית בירושלים, ואוכלוסיית הסקר תהיה משקי בית בירושלים). אוכלוסיית הסקר עלולה להיות "צרה" יותר לעומת אוכלוסיית המטרה, עקב מגבלות שונות (למשל, לא ניתן לאסוף מידע בחלק מהאזורים). ההבחנה בין אוכלוסיית המטרה ובין אוכלוסיית הסקר אינה תמיד ברורה לצרכני המידע שמופק מהסקר וייתכן שיש צורך להדגיש את ההבדל ביניהן, אם הוא קיים. פער בין אוכלוסיית המטרה והאוכלוסייה הנחקרת עלול להתרחש גם במקרה שמפעילים מערכת ניבוי והסקה ממדגם לאוכלוסייה שונה מזו שממנה המדגם נלקח. למשל, במצב של "אוכלוסיית על" (super-population) חוקר בודק אפקטיביות של תכנית התערבות על מדגם מסוים, במטרה לנבא את הצלחת התכנית על כל מי שבעתיד יוכל להשתתף בה; דוגמה נוספת היא חקירת אוכלוסיית המדינה כעת, במטרה להסיק לגבי אוכלוסייה דומה לה בעתיד².

¹ להרחבה ראו:

Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2004). *Survey Methodology*. Hoboken, NJ. J. Wiley.

² Hartley, H. O., & Sielken Jr, R. L. (1975). A "super-population viewpoint" for finite population sampling. *Biometrics*, 31, 411-422.

מסגרת הדגימה

מסגרת הדגימה כוללת את כל היחידות של אוכלוסיית הסקר, וממנה נלקח המדגם. מסגרת הדגימה עשויה להיות מפת אזורים, תקופת זמן שבה התרחשו אירועים, או מסמכים בארכיון. לעתים קרובות מסגרת הדגימה אינה כוללת את כל יחידות האוכלוסייה (למשל, בתי ספר שאינם בפקוח ובמימון המדינה עלולים לא להיות ברשימת המוסדות של מערכת החינוך בארץ). במקרה שמסגרת הדגימה קטנה מאוכלוסיית הסקר, החוקר יכול לבחור אחת מבין שתי אפשרויות: 1. להגדיר מחדש את אוכלוסיית הסקר (כדי להגביר הלימה למסגרת), דבר שעלול לפגום ברלוונטיות של היסק על האוכלוסייה המקורית; 2. להשאיר את ההגדרה המקורית של אוכלוסיית הסקר ולהתייחס להטיית הכיסוי, דבר שעלול להוביל לחולשה במסקנות הסטטיסטיות.

במקרה שקיימים מספר מקורות מידע על האוכלוסייה (למשל, כתובות דואר בביטוח לאומי, מספרי טלפון ברשויות המס וכיו"ב), ניתן לשלב בין מספר מסגרות דגימה, כאשר חלק מהפרטים שבמסגרת עשויים להופיע באופן חופף בין המסגרות וחלק עשוי להיות ייחודי לכל מסגרת.

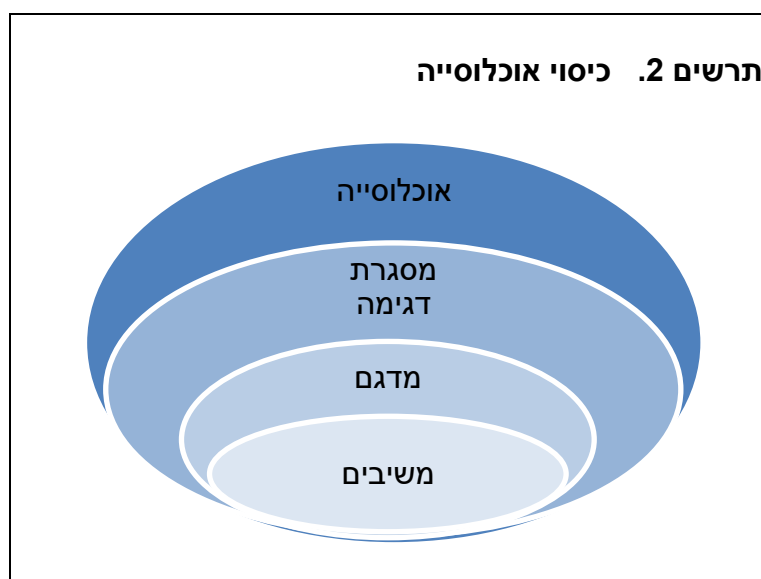
במקרה שלא קיימת מסגרת דגימה עבור כלל האוכלוסייה, ניתן לשלב דגימת אשכולות בשלבים ראשונים של הדגימה, ולהשתמש במסגרות דגימה בשלבים מאוחרים יותר של הדגימה (לדוגמה, בחירת אשכולות של אזורים סטטיסטיים בארץ ושל ישובים, ומתוכם דגימה אקראית או דגימת שכבות).

ייצוג

הטיית כיסוי

מסגרת הדגימה והמדגם עלולים שלא לייצג כיאות את האוכלוסייה. ראשית, אוכלוסיית המטרה עלולה להיות מכוסה באופן חלקי (under-coverage), אם חלק מיחידות האוכלוסייה אינן נכללות

במסגרת, או שהן נכללות במסגרת אך אינן נכללות במדגם (ראו תרשים 2). לחילופין, מסגרת הדגימה עלולה לכלול אלמנטים זרים, או יחידות פסולות שאינן חלק מאוכלוסיית הסקר (over-coverage), למשל, כאשר ברשימת משקי בית נכללים גם בתי עסק, שהם "מקרי אפס" (non-eligible units או ineligible units).



הטיות בדגימה

הדגימה עלולה להיות מוטית באופן שיטתי, כך שהיא לא תכלול חלק מהיחידות במסגרת הדגימה; מצב שלרוב ניתן להימנע ממנו כאשר לכל יחידה באוכלוסייה ישנו סיכוי **ידוע** להיכלל במדגם. בנוסף, המדגם עלול להיות לא מייצג, בייחוד כשקיימת שונות רבה באוכלוסייה; על כך ניתן להתגבר באמצעות מדגמים גדולים, שכוללים שכבות דגימה וללא אשכולות.

שיטות לדגימה הסתברותית

כאשר מטרת הסקר היא לספק אומדנים על משתנים מרכזיים באוכלוסיית המטרה, וכאשר קיימת מסגרת דגימה איכותית, רצוי להשתמש בשיטות הסתברותיות לדגימה. להלן מתוארים בקצרה ארבעה סוגים בסיסיים של דגימה הסתברותית.

דגימה אקראית פשוטה (Simple Random Sample)

בדגימה אקראית פשוטה לכל פרט יש סיכוי שווה להידגם. השונות בין יחידות במדגם מהווה אינדיקציה טובה לשונות באוכלוסייה, דבר שמקל על מתן אומדן לגבי דיוק התוצאות. שיטת דגימה זו עשויה להיות רגישה לטעויות בדגימה, כאשר המדגם קטן מאוד ו/או ישנן קבוצות ששיעורן באוכלוסייה נמוך ביותר. אם מטרת הסקר היא להשוות בין תת-אוכלוסיות, מדגם שכבות יתאים יותר (ראו בהמשך).

דגימה שיטתית (Systematic Sampling)

דגימה שיטתית נוחה לשימוש כאשר מסגרת הדגימה גדולה מאוד ו/או כאשר המדגם גדול מאוד. לאחר בחירה אקראית של הנדגם הראשון, שאר הנדגמים נבחרים לפי יחס נתון לנדגם הראשון (למשל, בחירה של כל נדגם 100 מתוך מסגרת דגימה של 30,000 תפיק 300 נדגמים). ניתן להתייחס לדגימה השיטתית כאל דגימה אקראית, כשסביר להניח שמסגרת הדגימה מסודרת בצורה מקרית ביחס למשתנים הנמדדים. אולם, ישנם מצבים שסידור הרשימה ו/או אופן הדגימה עלול לגרום להטיה. למשל, כשהרשימה היא מחזורית והמחזור חופף ליחס הדגימה – לדוגמה, בחירת כל בית דירות עשירי, כאשר כל הבתים הזוגיים פונים לחוף הים וכל הבתים האי-זוגיים פונים לצלע הר – דוגמה זו משקפת מצב של הערכת חסר של השונות באוכלוסייה.

דגימת שכבות (stratified sampling)

בדגימת שכבות, לפני עריכת הדגימה מחלקים את האוכלוסייה לתת-אוכלוסיות הומוגניות, כך שכל יחידה באוכלוסייה משתייכת לשכבה אחת בלבד. דגימת השכבות יכולה להיות פרופורציונלית לגודל השכבה באוכלוסייה או לא פרופורציונלית, למשל במקרה של קבוצות קטנות מאוד, אשר לאחר מעשה מקבלות תיקון למשקלן האמיתי באוכלוסייה. דגימת שכבות טובה לייצוג קבוצות מיעוט, כיוון שהיא מבטיחה ייצוג של המשתנים שאחראים לשכבות (לרוב משתני רקע), אשר קשורים גם למשתנים הנחקרים. אחד מחסרונותיה של דגימה זו, הוא שהמשתנה שאחראי לשכבה צריך להיות

מאופיין במסגרת הדגימה וכי ככל שמשתמשים במשתנים רבים יותר ליצירת שכבות כך המדגם בתוך כל שכבה קטן והאפשרות לטעות באומדנים גדלה. בנוסף, דגימה זו מגבילה את השימוש בשיטות מורכבות לניתוח סטטיסטי, אשר מניחות דגימה אקראית פשוטה (למשל, רגרסיה, ניתוח שונות, מבחנים א-פרמטריים).

דגימת אשכולות (cluster sampling)

בניגוד לדגימת שכבות, שבה האוכלוסייה מחולקת לקבוצות הומוגניות (ושונות זו מזו), בדגימת אשכולות האוכלוסייה מחולקת לקבוצות הטרוגניות, כשההנחה היא שכל אשכול מייצג את כלל האוכלוסייה. לאחר החלוקה לאשכולות נעשית דגימה אקראית של אשכול או של מספר אשכולות (דגימה בשלב אחד) וכל היחידות באשכול נחקרות; ניתן להמשיך ולדגום אקראית חלק מהיחידות באשכולות שנבחרו (דגימה בריבוי שלבים) ולחקור רק אותן. בנוסף, אם האשכולות נבדלים מאוד אלה מאלה בגודלם, ניתן לדגום מהם יחסית לגודלם. היתרון המרכזי של שיטת דגימה זו הוא שניתן להשתמש בה כשלא קיימת מסגרת דגימה עבור כלל האוכלוסייה. בנוסף, עלויות הסקר נמוכות יותר, לעומת שיטות דגימה אחרות. אחד מהחסרונות העיקריים של השיטה הוא שקיימת אפשרות שטעות הדגימה תהיה גדולה מהצפוי (למשל, אם האשכולות שנדגמו פחות הטרוגניים ולכן מייצגים פחות טוב את האוכלוסייה).

רקע

דגימה לא-הסתברותית קיימת בכל מצב שבו המדגם אינו נבחר בצורה אקראית (רנדומלית) מתוך האוכלוסייה, או שהסיכוי לבחירה של יחידה במדגם אינו ידוע.

השימוש בסקרי טלפון בחיגוי אקראי (Random Digit Dial) הרחיב את השימוש בדגימה הסתברותית לסקרי בחירות, סקרי שוק ומחקרים אקדמיים, אך העלייה החדה בשימוש בטלפונים סלולריים כאמצעי תקשורת יחיד של משקי הבית מטילה ספק במידת הכיסוי (coverage) של האוכלוסייה באמצעות סקרי הטלפון. בנוסף, חלה לירידה בשיעורי ההיענות של הציבור להשתתף בסקרים, אשר מעוררת את הצורך להתמודד עם נושא הטיית אי-ההשבה (non-response bias), הן בדגימה הסתברותית והן בדגימה לא-הסתברותית.

בשנים האחרונות השימוש בסקרים לא-הסתברותיים עלה מאוד, גם למטרות אמידה. במקרים רבים מדובר בסקרים אינטרנטיים, המבוססים על דגימה מתוך פנל. (ראו פירוט בפרק ב1 בחלק "דגימה מתוך פנלים"). השימוש המוגבר באינטרנט יצר ציפיות לגבי שימוש בסקרים מקוונים, אך ככלל,

³ להרחבה ראו:

Baker, R., Brick, J. M., Bates, N. A., Battaglia, M., Couper, M. P., Dever, J. A., ... Tourangeau, R., (2013a). *Report of the AAPOR Task Force on Non-probability Sampling. Technical report*, American Association for Public Opinion Research, Deerfield, IL.

Baker, R., Brick, J. M., Bates, N. A., Battaglia, M., Couper, M. P., Dever, J. A., ... Tourangeau, R., (2013b). Summary report of the AAPOR task force on non-probability sampling. *Journal of Survey Statistics and Methodology*, 1, 90–143.

Elliott, M. R., & Valliant, R. (2017). Inference for nonprobability samples. *Statistical Science*, 32(2), 249-264.

Mercer, A. W., Kreuter, F., Keeter, S., & Stuart, E. A. (2017). Theory and practice in nonprobability surveys: parallels between causal inference and survey inference. *Public Opinion Quarterly*, 81(S1), 250-271.

Pasek, J., & Krosnick, J. A. (2020). Relations Between Variables and Trends Over Time in RDD Telephone and Nonprobability Sample Internet Surveys. *Journal of Survey Statistics and Methodology*.

Scherpenzeel, A. C., and J. G. Bethlehem (2011). How representative are online panels? Problems of coverage and selection and possible solutions. In M. Das, P. Ester, & L. Kaczmarek (Eds.), *Social and behavioral research and the Internet: Advances in applied methods and research strategies* (pp. 105–132). New York: Routledge.

Vehovar, V., Toepoel, V., & Steinmetz, S. (2016). Non-probability sampling. *The Sage handbook of survey methods*, 329-345.

המחקר מראה כי מדגמים הסתברותיים מדויקים יותר, בנושאים רבים ומגוונים (למשל, הצבעה פוליטית, בריאות, צריכה, ומשתנים חברתיים-דמוגרפיים שונים). ישנם מחקרים שמראים כי למרות שמדגמים הסתברותיים מדויקים יותר בהפקת אומדנים נקודתיים (univariate estimates) ובהשוואות לאורך זמן (longitudinal analyses), אין הבדלים משמעותיים באומדנים לקשר בין משתנים (bivariate correlations) או בניתוחים רב-משתניים (multivariate analyses) (Cornesse et al., 2020)¹⁰.

עיקר ההבדל בין דגימה הסתברותית ללא-הסתברותית הוא סוג ההגנה בפני הסיכון לטעות. סקרים המבוססים על דגימה הסתברותית נותנים הגנה בפני הטיות כיסוי ואי-השבה באמצעות קביעת רמת מובהקות (לרוב, 5%) וחישוב של טווח ביטחון. לעומת זאת, סקרים שמבוססים על דגימה לא-הסתברותית אינם נסמכים על אותן ההנחות, כך שעליהם להתמודד עם בעיות כיסוי ואי-השבה באמצעות מודלים שונים.

חוקרים שמשתמשים בדגימה לא-הסתברותית נסמכים, לרוב, על אחד משני טיעונים אלה:

1. אין תלות בין התופעה שנבדקת להרכב המדגם – התופעה, או התהליך שנחקרים הם אוניברסליים, כך שהם מתרחשים אצל כלל האנשים. זו הנחה שמתקיימת ברוב המחקרים בתחום הפיזיולוגיה ובחלק מהמחקרים בפסיכולוגיה. בהתאם להנחה זו, צפוי שקבוצת אנשים מסוימת (המדגם) תנהג באופן דומה לזה של כל קבוצה אחרת (כל מדגם אחר), למעט במקרים שבהם ידוע על גורמים מתערבים. כלומר, ניתן לצפות שמשתני תוצאה מסוימים הם "עמידים" יותר ומושפעים פחות מדגימה לא-הסתברותית, כגון עמדות שנבדקות במסגרת סקר שוק כלפי מוצר זה או אחר. זאת, בשונה מבדיקת משתנים כלכליים-חברתיים מרכזיים כגון מעמד תעסוקתי, מצב בריאותי, מדדי עוני וכדומה.
2. הבדלים בין המדגם לאוכלוסייה יכולים לעתים לקבל מענה באמצעות התאמות – ניתן לערוך התאמות בדגימה (כגון matching) ובניתוח המידע (calibrations, adjustments), על מנת לדמות את הרכב המדגם לזה של האוכלוסייה. בהתאמות נעשה שימוש במשתנים מסוימים שמניחים כי הם משמשים כמקור להטיה פוטנציאלית. לפיכך, ההצדקה למידת הדיוק במדגמים לא-הסתברותיים, נובעת מהיכולת לזהות את כל המשתנים המתערבים ולשלוט בהם.

שיטות לדגימה לא-הסתברותית

בשונה משיטות הדגימה ההסתברותיות, אין תיאוריה או רציונל אחד, דומה, שנמצא בבסיס השיטות הלא-הסתברותיות. לשם הנוחות, מקובל לאפיין את הסוגים השונים של דגימה לא-הסתברותית לשלוש קבוצות עיקריות: דגימה מותאמת, דגימה באמצעות רשת קשרים ודגימת נוחות (Baker, et al., 2013a). היתרונות של דגימה לא-הסתברותית הם נוחות השימוש, קיצור זמן האיסוף וחיסכון

יחסי בעלות. החיסרון העיקרי של דגימה לא-הסתברותית הוא הקושי להעריך את הדיוק ואת הטעות באופן כמותי.

הבחירה בין דגימה מותאמת, דגימה באמצעות רשת קשרים או דגימת נוחות צריכה להיות בהתאם למטרות הסקר. דגימה מותאמת מאפשרת לתת הערכות על כלל האוכלוסייה, בעוד שדגימת רשת קשרים מיועדת לחקירה של אוכלוסייה קטנה או לא נגישה.

דגימה מותאמת (sample matching)

לעומת שיטות לא-הסתברותיות אחרות, דגימה מותאמת מאפשרת לעצב (design) את המדגם, כך שהוא ישקף מאפיינים מרכזיים של אוכלוסיית המטרה. הדבר נעשה על ידי השוואה לנתוני מפקד, נתונים מנהליים או מדגם הסתברותי. לרוב מדובר במאפיינים נגישים, כגון גיל ומגדר, אך בהתאם למטרות הסקר, רצוי לכלול מגוון רחב של משתנים רלוונטיים כדי לאפשר השוואה בין הנדגמים בסקר לאוכלוסיית המטרה, כולל משתני התנהגות (כגון עיסוק בספורט) ועמדה (למשל, דעה פוליטית). המודל שקובע את עיצוב המדגם אמור להגביר את הדיוק של האומדנים שיופקו ממנו, באמצעות הכללת משתני עזר (auxiliary variables), שמטרתם לזהות מקורות אפשריים להטיה – הטיית דגימה והטיית אי-השבה (ראו "שיטות להתאמה ותיקון בשלב הדגימה").

במקור, דגימה מותאמת שימשה במערך מחקר דמוי ניסוי, להתאמה בין קבוצת ניסוי וקבוצת ביקורת באמצעות משתני בקרה (covariates), או משתני עזר, כשההנחה היא שמשתנים אלה קשורים הן למשתנים המסבירים (בלתי תלויים) והן לתוצאות (המשתנים התלויים). השימוש בדגימה מותאמת לעריכת סקרים התפתח מאוד במהלך השנים האחרונות.

דגימת מכסה (quota sampling) תיחשב סוג של דגימה מותאמת, אם היא כוללת מכסות כלכליות-חברתיות אשר משקפות התפלגויות באוכלוסייה של משתנים שהם משמעותיים לסקר. דוגמאות למשתנים כאלה הן למשל, אזור גיאוגרפי, מגדר, גיל, עמדה פוליטית, ועוד. היחס בין קבוצות באוכלוסייה (כגון נשים וגברים) יכול להיות זהה במדגם, או בשיעור שונה שמתוקן לאחר מכן.

אחת הדרכים הנפוצות לעריכת דגימה מותאמת היא באמצעות פנל לא-הסתברותי. המאפיין המרכזי של פנל לא-הסתברותי הוא שהוא מורכב מאנשים שפנו להצטרף אליו. לרוב, הפונים להצטרף לפנל מתבקשים למלא שאלון פרופיל, ובו נאספים משתני עזר שונים, המאפשרים לערוך דגימה מותאמת. מבין המשתתפים בפנל נבחרים מעת לעת משיבים לסקרים שונים (עבור תשלום או קבלת הטבות עם שווי כספי) במשך תקופה מסוימת. במובן זה, סוקרים רבים מתייחסים לפנל כמעין "מסגרת דגימה". (ראו הרחבה בהמשך בחלק "דגימה מתוך פנלים").

דגימה באמצעות רשת קשרים (network sampling)

קיימות אוכלוסיות המתאפיינות בין השאר על ידי מערכת קשרים בין הפרטים. מערכת כזו יכולה לשמש גם מושא של מחקר וגם אפיק (link-tracing design) להגיע לפרטים באוכלוסייה שהנגישות

אליהם מוגבלת, כגון בחקירה של אוכלוסייה נדירה (פחות מעשירית האחוז מהאוכלוסייה הכללית)⁴ או חבוייה, שאין עבורה מסגרת דגימה (למשל, נוער דר רחוב, משתמשים כבדים בסמים, אנשים שאוהבים אוזו, קבוצות חברתיות באינטרנט). אם ניתן לאתר מספר משתתפים וליצור איתם אמון, הם עשויים לקשר את החוקרים לחברים נוספים ב"רשת".

אחת השיטות היא "התרחבות הקשרים" (network scale-up)⁵, שבה משיבים במדגם מייצג של האוכלוסייה הכללית נשאלים על מספר האנשים שהם מכירים מהאוכלוסייה הכללית ומספר האנשים שהם מכירים מאוכלוסיית המטרה (למשל, בעלי כלבים); הכפלת גודלה של האוכלוסייה הכללית ביחס שבין הקשרים שיש לאדם מסוים עם אנשים מאוכלוסיית המטרה ובין סך כל הקשרים שיש לו באוכלוסייה הכללית נותן אומדן לגודלה של אוכלוסיית המטרה. היתרון בשיטה הוא שהחוקר אינו נדרש לגשת ישירות לאוכלוסיית המטרה והוא יכול לשמור על פרטיות המשיבים כיוון שהוא אינו אוסף מידע אישי עליהם. כדי לצמצם הטיות באומדנים על החוקרים להתמודד עם קשיים במדידה, כגון אלה: יש משיבים שאינם מודעים להשתייכות של חלק ממכריהם לאוכלוסיית המטרה (מידע רגיש לרוב מוסתר על ידי אוכלוסיית המטרה); לאנשים באוכלוסייה הכללית יש סיכוי שונה ליצירת קשרים עם אנשים מאוכלוסיית המטרה (למשל, עקב מאפיינים דמוגרפיים של המשיב ו/או חסמים גיאוגרפיים); כיוון שבסקר ניתן מעט זמן להיזכרות יש נטייה לתת הערכות יתר לגודלן של קבוצות קטנות ולתת הערכות חסר לגודלן של קבוצות גדולות; עלולות להתקבל תשובות מוטות במכוון כדי להסתיר היכרות של המשיב עם אנשים מקבוצות דחיות חברתית.

שיטה נוספת היא "דגימה באמצעות נדגמים" (RDS - respondent-driven sampling)⁶; ניתן לכנות אותה דגימת "כדור שלג עם מתודולוגיה". השיטה פותחה על מנת לתת מענה להטיות בדגימה לא-הסתברותית ולאפשר לחקור אוכלוסיות חבויות או קשות להשגה, אשר אין עבורן מסגרת דגימה וישנה רגישות בחקירתן ולכן קושי בקבלת שיתוף פעולה (למשל, משתמשים בסמים). בשלב ראשון פונים למספר קטן של חברים מהקבוצה, המשמשים כ"זרעים" למדגם וכל משתתף נשאל על מספר החברים שהוא מכיר בקבוצה, והוא מתבקש לגייס שלושה מהם לסקר. כל משתתף מקבל תגמול על השתתפותו (באמצעות קופון הטבה המשמש לזיהויו בסקר) וכן על כל משתתף שהוא מגייס. לרוב, מושג מדגם גדול דיו שהינו בלתי תלוי במאפייני המגויסים הראשונים, לאחר כשישה גלי גיוס. לשיטה

-
- ⁴ Kish, L. (1987). *Statistical Design for Research*. New York: John Wiley & Sons, Inc.
- ⁵ Laga, I., Bao, L., & Niu, X. (2020). Thirty Years of The Network Scale up Method. *arXiv preprint arXiv:2011.12516*.
- ⁶ Heckathorn, D. (1997). Respondent driven sampling: A new approach to the study of hidden samples. *Social Problems*, 44, 174–199.

זו יתרון באיתור יעיל של נדגמים, בכך שאין מסירת פרטים מזהים על ידי גורם שלישי, ובשימוש בהנחה כי רשת הקשרים מאפשרת לפרטים באוכלוסייה סיכוי גדול מאפס להיכלל במדגם.

שיטה אחרת לאמידה של אוכלוסייה חבויה היא באמצעות איתור כל המקומות והזמנים שבהם מרוכזים חברי הקבוצה ("venue-based method" or "time-location sampling") ודגימה הסתברותית מתוך מסגרת זו. לשימוש בשיטה כזו יש להניח כי ניתן לאתר את כל, או לפחות רוב, המקומות והזמנים שבהם חברי הקבוצה מתכנסים, כי שיעור גבוה מחברי הקבוצה אכן מבקרים במקומות וזמנים אלה, וכי יש יציבות במקומות ובזמני הכינוס של חברי הקבוצה.⁷

דגימת נוחות (convenience sampling)

בדגימת נוחות השיקול העיקרי הוא נוחות האיתור או גיוס המשתתפים לסקר. שיקולים נוספים יכולים לכלול גם חיסכון בעלויות. מדגם נוחות עשוי לכלול קבוצת אנשים שנמצאים במקרה בקרבת אוסף המידע, כגון אנשים שמחכים בתחנת אוטובוס, בקניון (כדוגמת סקרי יירוט במרכזי קניות - mall intercepts), או "קהל שבוי" כמו סטודנטים בשיעור של מרצה. פנייה למשתתפים במדגמי הנוחות יכולה להיעשות באמצעים שונים; למשל, באמצעות הפצת עלונים במקומות ציבוריים (מלונות, מסעדות, משרדים), פנייה להביע השקפות ודעות באמצעות SMS לטלפון הנייד על נושאים שקשורים לבידור (תחרות שירים בטלוויזיה) או לפוליטיקה (סקרי בחירות), פנייה להביע מידת שביעות רצון לאחר קנייה בחנות או קבלת שירות, דרך הטלפון הנייד או פנים אל פנים, וכדומה. כיום, פנייה להשתתף בסקרים נפוצה מאוד דרך האינטרנט, למשל באמצעות דייג (river sampling). במצב זה, לרוב אין איסוף מראש של מאפיינים דמוגרפיים ואחרים של המשיב, כך שלא ניתן "לעצב" את המדגם בהתאם להתפלגויות ידועות של משתנים רלוונטיים באוכלוסייה. יש לציין שמדגם מכסה שאינו סוג של דגימה מותאמת הוא סוג של סקר נוחות.

מדגם כדור שלג (או "דגימת שרשרת"), אשר אוסף משיבים באופן מתגלגל, נחשב כמדגם נוחות. שיטת דגימה זו נועדה להגיע מהר לגודל מדגם רצוי, באמצעות "גלגול" השאלונים לקבוצות גדולות ככל האפשר. שיטה זו שונה מ"דגימה באמצעות נדגמים" (ראו לעיל), שהיא מבוססת על "גלגולים" רבים לקבוצות קטנות, במטרה להגיע לבסוף למדגם מייצג.

בקבוצה זו של מדגמי נוחות ניתן לציין סוגים נוספים כגון דגימה מכוונת או שיפוטית (purposive sampling or judgmental sampling), שבה מומחים בוחרים את המקרים הטיפוסיים או החריגים (למשל, מקרים של הצלחות לעומת מקרים של כישלונות), באמצעותם אפשר לאפיין וללמוד לעומק את התופעה הנחקרת. "חקר מקרה" (case study) הוא דוגמה קיצונית ללמידה על תופעה באמצעות יחידה אחת או מספר יחידות בודדות. סוגי מדגמים אלה משמשים בעיקר לבדיקת השערות

⁷ Magnani, R., Sabin, K., Saidel, T., & Heckathorn, D. (2005). Review of sampling hard-to-reach and hidden populations for HIV surveillance. *Aids*, 19(suppl 2), S67-S72.

או לחקירה מעמיקה של תופעות וברוב המקרים הם אינם מתאימים לחקירה כמותית, המבוצעת באמצעות סקר.

ככלל, השימוש במדגמי נוחות אינו מתאים לשם הסקת מסקנות על האוכלוסייה. כיוון שלא מדובר בדגימה הסתברותית, קשה לדעת אם המדגם מכסה את כל האוכלוסייה הרלוונטית, ובנוסף, הבחירה העצמית למדגם עלולה להוביל לריבוי של מקרים יוצאי דופן (outliers). אם בכל זאת נעשה שימוש בדגימת נוחות, יש לפרט את ההבדלים בין המדגם שנעשה בו שימוש לעומת מדגם אקראי ולציין אם יש אפשרות שקבוצות מסוימות לא קיבלו ייצוג במדגם או שיש כאלה שקיבלו ייצוג יתר. כפי שנאמר, לא ניתן להסיק מהמדגם לכלל האוכלוסייה, אולם ניתן לומר שקשרים מסוימים, אם נמצאו בין משתנים, עשויים להימצא גם באוכלוסייה⁸.

הטיות ייצוג בשיטות לדגימה לא-הסתברותית

כאמור, הטיית כיסוי, שהינה סוג של הטיה בייצוג האוכלוסייה, מתרחשת כאשר מסגרת הדגימה אינה מכסה באופן מלא את האוכלוסייה. הטיית כיסוי במדגמים לא-הסתברותיים יכולה לנבוע מזה שחלק מהאוכלוסייה אינו מכוסה או מזה שיש לו פוטנציאל נמוך (או גבוה) יחסית להיות מכוסה (Elliott and Valliant, 2017), כגון קשישים לעומת צעירים בפנלים אינטרנטיים. ניתן להמחיש הטיית כיסוי בדוגמה על סקר אינטרנטי שנערך על אוכלוסיית בוגרים מעשנים. חלק מהאוכלוסייה שאין לה גישה לאינטרנט היא אוכלוסייה שלא מכוסה בסקר; בנוסף, ישנה אוכלוסייה שיש לה אפשרות לכיסוי אשר כוללת אנשים שיש להם גישה לאינטרנט, אך מבקרים באתרים שלא נמצאים ברשימת האתרים בהם נערך הגיוס לסקר; החלק מהאוכלוסייה המכוסה הוא זה שיש לו גישה לאינטרנט ושמבקר באתרים שבהם מתרחש גיוס המתנדבים. לסיכום, המדגם כולל רק את מי שיש לו אפשרות להשתתף בסקר ואשר נענה לכך. יש לשים לב כי החלק מתוך האוכלוסייה שאינו מכוסה כולל אנשים שאין להם גישה לאינטרנט ואנשים שיש להם גישה לאינטרנט אך אינם מבקרים באתרים שבהם מתרחש הגיוס לסקר. במצבים רבים, החלק מתוך האוכלוסייה שאינו מכוסה עלול להיות גדול יותר מהחלק המכוסה (Elliott and Valliant, 2017).

שיטות להתאמה ותיקון בשלב הדגימה

חוקרים משתמשים בשיטות עקיפות ובשיטות ישירות, על מנת לדמות תוצאה של שיטת דגימה הסתברותית. במידת האפשר, יש לשלב מאפיינים של דגימה הסתברותית באופן ישיר. למשל, ניתן לשלב דגימה הסתברותית ממסגרת דגימה בשלב הראשון ודגימת נוחות בשלב השני (לדוגמה, דגימה אקראית של בתי ספר מרשומות של משרד החינוך, פנייה לתלמידים בבתי ספר אלה

⁸ Rivera, J. D. (2019). When attaining the best sample is out of reach: Nonprobability alternatives when engaging in public administration research. *Journal of Public Affairs Education*, 25(3), 314-342.

להשתתף במדגם באמצעות מודעות בבית הספר ובאתר האינטרנט של כל בית ספר). כמו כן, ניתן לשלב את עקרון האקראיות בכל שלב בדגימה לא-הסתברותית, למשל, בסקרי יירוט במרכזי קניות, ניתן להחליט שפונים לכל לקוח X שמגיע לאזור מסוים, או לפנות ללקוח אחרי כל Y שניות וכדומה.

אחת מהדרכים העקיפות שניתן להשתמש בה כדי להוסיף מאפיינים של דגימה הסתברותית היא לכסות חלקים רבים ככל האפשר של האוכלוסייה, באמצעות פריסה רחבה של המדגם; למשל, באמצעות פרסום מודעות לגיוס המשיבים במגוון אתרים באינטרנט, בצורות מדיה שונות, באזורים שונים בארץ ועוד.

בשיטה נוספת נעשה שימוש במשתני עזר (auxiliary variables)⁹ לעתים קרובות על מנת להתגבר על הטיית אי-השבה. משתנה עזר טוב הוא גורם משותף שקשור הן למשתנה הנחקר והן לנטייה להשיב (או לא להשיב). משתני עזר אינם מהווים חלק ממודל הניתוח (כך למשל, לא כדאי לקחת כמשתני עזר משתנים שאחראים לריבוד השכבות), אלא יש להתייחס אליהם כמשתנים מנבאים מכיוון שהם יכולים לשמש לזקיפה של ערכים חסרי תצפיות.

ישנם מספר תנאים לשימוש במשתנה עזר, אך לרוב לא ניתן לעמוד בכולם: ראשית, ערכי המשתנה צריכים להיות ידועים עבור המשיבים ועבור הלא משיבים; שנית, יש למדוד את המשתנה באופן מלא וללא הטיה עבור כל היחידות שנדגמו; שלישית, משתנה העזר רצוי שיהיה במתאם גבוה (רצוי $r = .70$ ומעלה) עם המשתנים שנחקרים; רביעית, משתנה העזר רצוי שיהיה במתאם גבוה גם עם הנטייה להשיב לסקר, ובכך להפחית את הטיית אי השבה.

התאמת ההתפלגויות של משתני העזר במדגם יוצרת למעשה מדגם מותאם, במטרה לשפר את אופן הייצוגיות של האומדנים לפרמטרים באוכלוסייה. השוואת ההתפלגויות בין המדגם לאוכלוסייה יכולה להיות ברמת הפרט או ברמת הקבוצה. התאמה של "אחד לאחד" נעשית באמצעות שימוש במדגם הסתברותי איכותי קיים, כאשר הנדגמים במדגם הלא-הסתברותי נבחרים בהתאם למאפייני הנדגמים במדגם ההסתברותי (על פי אותם משתני עזר שנקבעו). תגובות המשיבים במדגם הלא-הסתברותי משמשות לאחר מכן להפקת אומדנים, לאחר שקלול נוסף שמתחשב באי-השבה. כיוון שריבוי של משתני עזר מקשה על התאמה של "אחד לאחד", דרך שכיחה היא לערוך התאמה ברמת הקבוצה. בהתאמה מסוג זה נערכת השוואה של התפלגות משתני הבקרה בין שני המדגמים, או בין המדגם לכלל האוכלוסייה (למשל לפי נתוני המפקד, הסקר החברתי, סקר כוח אדם וכדומה). אם רוצים להתגבר על ריבוי של משתני עזר, ניתן לחשב "ציוני נטייה" (propensity scores), ברמת היחידה הנדגמת. הפרוצדורה הבסיסית כוללת התאמה של יחידות במדגם הלא-הסתברותי עם יחידות במדגם הסתברותי גדול, בהתאם למשתנים זהים שנאספים בשני המדגמים. לרוב עורכים ניתוח של

⁹ Olson, K. (2013). Paradata for nonresponse adjustment. *The Annals of the American Academy of Political and Social Science*, 645(1), 142-170.

רגרסיה לוגיסטית שבה משתני העזר מנבאים את ההסתברות לכך שנדגם ישתתף או לא ישתתף במדגם הלא-הסתברותי.

התקווה שעומדת בבסיס של מדגם מותאם היא שהמשתנים שמשמשים לקביעת הקבוצות במדגם מנטרלים את כל המקורות האפשריים להטיה. אם דבר זה מושג, ההיסקים שייעשו על בסיס המדגם יהיו מדויקים. בניגוד לדגימה לא-הסתברותית, אחד מהיתרונות של דגימה אקראית הוא, שהיא מפיקה איזון אוטומטי של משתני הבקרה, גם אם ההתאמה שלהם אינה מושלמת. בדגימה שאינה הסתברותית אם אין בחירה נכונה של משתני הבקרה ו/או הם אינם זמינים, הטיות הכיסוי והבחירה תישארנה. בנוסף, כיוון שברוב המקרים בודקים הרבה משתני תוצאה בסקר אחד, יש צורך לאתר ולהתאים משתני בקרה רבים ושונים, או להשתמש בציוני נטייה. האתגר בשימוש בדגימה מותאמת הוא שהדיוק של המסקנות יכול להיבחן רק באופן אמפירי, לאחר מעשה ובאופן מצטבר (על פני בדיקות). בנוסף, מכיוון שבחירת המשתנים במודל הדגימה נעשית בהתאם לשאלת המחקר ולאופי הניתוחים המבוקשים, אין ערבות לכך שניתן לענות על שאלות שונות (ולתת אומדנים שונים) ממדגם בעל הרכב נתון, שעוצב במטרה לענות על שאלה מסוימת (ולספק אומדנים מסוימים).

דגימה מתוך פנלים¹⁰

לעתים הסקר נערך באמצעות משתתפים בפנל. פנלים, לרבות פנלים אינטרנטיים, כוללים אנשים שהסכימו להשתתף או שפנו להשתתף בסקרים לאורך זמן. הפנל יכול להיות הסתברותי או לא-הסתברותי ודרך איסוף הנתונים עשויה להיות באמצעות האינטרנט, הטלפון, הדואר, פנים אל פנים, או כל שילוב של דרכים, בהתאם למטרות הסקר. בשנים האחרונות הוקמו פנלים אינטרנטיים על ידי ארגונים שונים ברחבי העולם, רובם המכריע אינו הסתברותי. עם זאת, גופי ממשל ומחקר במדינות רבות עושים מאמץ לפיתוח של פנלים אינטרנטיים הסתברותיים, כגון - American Trends Panel - ATP של ארגון PEW, AmeriSpeak panel של מרכז המחקר הלאומי לדעות (NORC) באוניברסיטת שיקגו, LISS - Longitudinal Internet Studies for the Social Sciences ההולנדי, German Social Science Infrastructure Services - GIP German Internet Panel - ו- GESIS הגרמנים, NatCen Panel הבריטי, ועוד.

פנלים הסתברותיים מופקים מדגימה אקראית מתוך האוכלוסייה באמצעות מסגרת דגימה, למשל מרשם התושבים, רשימת כתובות דואר (בארצות הברית) וכדומה. הדרך השגורה במרבית הפנלים

¹⁰ להרחבה ולהמלצות ראו:

Cornesse, C., Blom, A. G., Dutwin, D., Krosnick, J. A., De Leeuw, E. D., Legleye, S., ... Struminskaya, B. (2020). A Review of Conceptual Approaches and Empirical Evidence on Probability and Nonprobability Sample Survey Research. *Journal of Survey Statistics and Methodology*, 8, 4–36.

Callegaro, M., & DiSogra, C. (2008). Computing response metrics for online panels. *Public opinion quarterly*, 72(5), 1008-1032.

ההסתברותיים, היא לפנות לנדגמים, להסביר להם על הפנל ולבקש את הסכמתם להשתתף בו. לאחר זאת, לרוב נשלח למסכימים שאלון פרופיל, שכולל שאלות שונות לאפיון של המשיב (כגון נתונים דמוגרפיים, עמדות בנושאים שונים ועוד). פנלים לא-הסתברותיים (למשל, opt-in panels או access panels) מבוססים על מתנדבים שפנו לפנל. בפנלים כאלה מסגרת הדגימה אינה ידועה (כלומר, אותה מסגרת שבה רשומה לכאורה כל האוכלוסייה שמשתמשת באינטרנט וגולשת באתרים שבהם פורסמה האפשרות להצטרף לפנל). בפנלים שאינם הסתברותיים, ניתן להצטרף באופן מיידי לפנל או לאחר שמנהל הפנל מוודא כי כתובת הדואר האלקטרוני היא אכן של הפונה. לבידור זה יש עדיפות, כיוון שכך ניתן למנוע ריבוי השבה על סקר על ידי אותו משיב. אם המשיב ממלא שאלון פרופיל לפני ההשתתפות בסקר, ניתן לערוך דגימה מותאמת מתוך הפנל (ראו פירוט לעיל). אם אין אפשרות לערוך דגימה מותאמת, תהיה התוצאה מדגם שאינו מייצג את האוכלוסייה.

אחד ממדדי האיכות של סקר שנעשה באמצעות פנל אינטרנטי הוא באמצעות חישוב של שיעורי ההשתתפות. בשלב ראשון בפנל הסתברותי מחשבים את שיעור המסכימים להשתתף בפנל מתוך כל מי שנדגם ממסגרת הדגימה (שיעור הגיוס – Recruitment rate). בשלב שני בפנל הסתברותי (ראשון בפנל לא-הסתברותי) מחשבים את השיעור מצטרפים שמלאו שאלון פרופיל עם פרטי רקע (שיעור השלמת פרופיל – Profile rate). לאחר עריכת הסקר, מחשבים את השיעור של חברי הפנל שהשלימו מענה לשאלון (שיעור השלמת השאלון – Completion rate). לבסוף, שיעור השבה כולל (cumulative response rate) מחושב כמכפלה של שיעור הגיוס, השלמת שאלון הפרופיל והשלמת שאלון הסקר; אם מתקיימים מספר גלי בדיקה, נוסף לחישוב גם שיעור התמדה (retention rate) - מספר המשיבים בגל עוקב מתוך מספר המשיבים בגל הקודם). בפנלים אינטרנטיים לא-הסתברותיים לא ניתן לחשב שיעור השבה כולל כמתואר, כיוון שלא ניתן לחשב את שיעור הגיוס (כיוון שהדגימה לא נעשתה מתוך מסגרת דגימה). עם זאת, חשוב לבחון את שיעור השלמת השאלון, שמשקף את העניין שהסקר מצליח לעורר במשיבים ואת שיתוף הפעולה שלהם וכן את שיעור הנשירה מהפנל, שכאשר הוא גבוה מאוד הוא עשוי לאותת על קושי בסקרים (למשל: סקר ארוך ו/או שאלות קשות להבנה וכדומה).

המושג של "חברות פעילה" דומה לכאורה בין פנלים שונים, כיוון שמדובר בכל מי שנמצא בפנל לאחר הגיוס ופנוי למלא סקר. אולם בפועל, בעיקר בפנלים לא-הסתברותיים, הפעילות בפנל תלויה בגורמים שונים. ראשית, היא תלויה במשתתף, שלעיתים בוחר שלא לענות על סקר, או שהוא אינו יכול לענות (למשל: עקב מחלה או חופשה). בנוסף, מידת הפעילות בפנל תלויה גם במדיניות של הפנל, שעשויה להיות נוקשה יותר או רכה יותר לגבי מספר הסקרים ו/או פרקי הזמן שבהם מותרת השתתפות ואי-השתתפות. בנוסף, חבר בפנל יכול להיות "לא פעיל", אם חוקר החליט לבחור רק חברי פנל עם מאפיינים מסוימים, או שמנהל הפנל מעוניין להפחית את הנטל ממשיבים שמשתתפים בסקרים רבים. כמו כן, יש שונות בין הפנלים לגבי הגבלת משך פעילות אפשרי של חברי הפנל ולגבי תדירות הריענון של הפנל. הנושא של "נשירה" (attrition) קשור לחברות הפעילה, וגם הוא משתנה בין פנלים.

בפנלים אינטרנטיים לא-הסתברותיים נשירה מוגדרת למעשה כשיעור המשתתפים שנשרו מהפנל בתקופה מוגדרת. בפנל אינטרנטי הסתברותי, המושג של נשירה מתייחס לשיעור המשיבים בגל (cohort) נתון יחסית לגל קודם, ויש לו משמעות רבה לגבי ייצוגיות של קבוצות מסוימות באוכלוסייה ולתוספת בחישוב של שיעור השבה כולל בסקרים חוזרים ובסקרי אורך (longitudinal).

תכנון כלי מדידה ושיטות איסוף

כדי להבטיח איכות של הסקר, לא רק שיש להוכיח שהמדידה לא תושפע מהטיות ייצוג (כיסוי אוכלוסייה, חוסר הטיה בדגימה ובהשבה), יש לדאוג שהיא תהיה תקפה, מהימנה וללא הטיות מדידה.

תוקף ומהימנות של המדידה¹¹

תוקף ומהימנות חשובים לבדיקת איכות המדידה של היבטים אובייקטיביים (כגון: משקל, גיל, השכלה) ולא אובייקטיביים (כגון: עמדות, תפיסות, מאפייני אישיות). תוקף משקף את המידה שבה כלי המדידה מודד את מה שהוא אמור למדוד; מהימנות משקפת את מידת העקביות או היציבות של המדידה.

נהוג לבטא תוקף של כלי מדידה על ידי מספר מדדים עיקריים, כפי שיפורט בהמשך: תוקף תוכן (content validity), שנכלל בשלבים ראשונים של פיתוח מדד לבדיקת נושא החקירה וכן תוקף מבנה (construct validity) ותוקף תלוי קריטריון (criterion related validity), אשר מאמתים את תוקף הכלי שנבנה בשלב קודם באמצעות השוואה למדדים אחרים. מהימנות של המדידה נבדקת לרוב באמצעות מהימנות פנימית (internal reliability) ומהימנות מבחן חוזר (test re-test reliability).

¹¹ להרחבה ראו:

Alwin, D. F. (2016). Survey data quality and measurement precision. In C. Wolf, D. Joye, T. Smith, and Y. Fu (Eds.), *The SAGE handbook of survey methodology* (pp. 527-557). London: SAGE Publications Ltd.

Almanasreh, E., Moles, R., & Chen, T. F. (2019). Evaluation of methods used for estimating content validity. *Research in social and administrative pharmacy*, 15(2), 214-221.

Bannigan, K., & Watson, R. (2009). Reliability and validity in a nutshell. *Journal of clinical nursing*, 18(23), 3237-3243.

Clifton, J. D. (2020). Managing validity versus reliability trade-offs in scale-building decisions. *Psychological Methods*, 25(3), 259.

Drost, E. A. (2011). Validity and reliability in social science research. *Education Research and perspectives*, 38(1), 105-123.

Hughes, D. J. (2018). Psychometric validity: Establishing the accuracy and appropriateness of psychometric measures. In P. Irwing, T. Booth, & D. J. Hughes (Eds.), *The Wiley handbook of psychometric testing: A multidisciplinary approach to survey, scale and test development*. Chichester, UK: Wiley, 751-779.

Ponterotto, J. G., & Ruckdeschel, D. E. (2007). An overview of coefficient alpha and a reliability matrix for estimating adequacy of internal consistency coefficients with psychological research measures. *Perceptual and motor skills*, 105(3), 997-1014.

תוקף תוכן (content validity)

תוקף התוכן משקף את מידת הרלוונטיות והייצוג של התחומים שנועדו להימדד, על ידי פריטי השאלון. במילים אחרות: תוקף תוכן הוא המידה שבה שאלון, או מדד (אינדקס) בשאלון, מודד את תחומי התוכן שהוא אמור למדוד. ביתר פירוט, תוקף התוכן נוגע לדרך שבה הנושא שנבחר להימדד מתבטא בהגדרה האופרציונלית; למידת הייצוגיות של תחום התוכן שהוגדר למדידה; למידת הרלוונטיות של רכיבי השאלון לתחום שהוגדר (כך שלא ייצגו תחומים אחרים) ולמידת ההתאמה של שיטות הפיתוח שנקטו להשגת הייצוגיות והרלוונטיות של הרכיבים. שני האיומים המרכזיים על תוקף תוכן הם תת-ייצוג של תכנים רלוונטיים וייצוג של תכנים לא רלוונטיים על ידי רכיבי השאלון. ללא התמודדות מתאימה עם איומים אלה תהיה פגיעה במהימנות הפנימית של השאלון (ראו בהמשך).

את בדיקת תוקף התוכן יש לבצע בשלבים הראשונים של פיתוח השאלון. תהליך תיקוף התוכן כולל שלושה שלבים: פיתוח, שיפוט והתאמה. שלב הפיתוח כולל הגדרה מושגית ואופרציונלית של התחום שנבדק ולאחר מיכן הפקת פריטים, תוך שימוש נרחב בכל מקור מידע רלוונטי לרבות סקירת ספרות, ראיונות עם מומחים ובמידת הצורך ראיונות עם נציגים רלוונטיים מהאוכלוסייה. שלב השיפוט נעשה לאחר שהפריטים פותחו וסודרו. קבוצת מומחים (בין חמישה לעשרה) בוחנת את מידת הרלוונטיות והייצוגיות של כל פריט בשאלון וכן את ייחודיותו, חשיבותו ובהירות ההגדרה האופרציונלית שהוא משקף. יש שיטות שונות לחישוב התוקף של המדד כולו ושל כל פריט בנפרד; רוב השיטות מפיקות ערך של יחס שמתחשב במספר השופטים ובמספר הפריטים ו/או בעוצמת הדירוג שקיבל כל פריט. בשלב האחרון מחליטים לגבי כל פריט אם להשאירו, לשנותו או להוציאו מהשאלון וכן אם יש צורך בפיתוח של פריטים חדשים, דבר שדורש חזרה על שלבי התיקוף עד להסכמה על כל פריטי המדד.

תוקף מבנה (construct validity)

תוקף מבנה משקף את המידה שבה כלי המדידה שבו משתמשים מצליח לתרגם תפיסה תיאורטית של מושג, רעיון או התנהגות לרכיבים אופרציונליים, מציאותיים. תוקף המבנה מבטא את מידת ההלימה של התפיסה התיאורטית עם המדד האופרציונלי תוך שימוש בגישה לא ישירה להוכחת הקשר. ניתן למדוד את תוקף המבנה באמצעות תוקף מתכנס (convergent), מבחין (discriminant) וגורמי (factorial) ולפי העניין גם בתפקוד שונה בין קבוצות ידועות (למשל: בריאים לעומת חולים במחלה מסוימת).

תוקף מתכנס מדגים את ההלימה בין שני מדדים שונים שאמורים לשקף את אותו מבנה תיאורטי. ככל שהמדידה תדגים מתאם גבוה יותר בין שני המדדים, כך הוודאות לגבי תוקף המבנה עולה. למשל: בדיקת מתאם בין שאלון חדש שפותח לבדיקת חרדה ובין שאלון קיים שמשמש לבדיקת חרדה (בדיקה דומה למדידת רגישות - sensitivity). מנגד, תוקף מבחין מבטא חוסר מתאם (או מתאם נמוך מאוד) בין שני מדדים שאמורים לשקף מבנים תיאורטיים שונים. ככל שהמדידה תדגים מתאם נמוך יותר כך הוודאות לגבי תוקף המבנה עולה. למשל: בדיקת מתאם בין שאלון חדש שפותח

לבדיקת חרדה ובין שאלון קיים שמשמש לבדיקת אפיון קרוב, אך שונה, כגון דיכאון (בדיקה דומה למדידת סגוליות - specificity).

תוקף מתכנס ומבחין קשורים לגישת "ריבוי תכונות ושיטות" (multitrait-multimethod - MTMM). בגישה זו נעשית בדיקה בו-זמנית למספר נושאים, או תכונות, באמצעות מספר כלים, או שיטות, כאשר הציפיה דומה למתואר לעיל: מתאמים גבוהים בין אותה תכונה שנמדדת באמצעות כלים שונים ומתאמים נמוכים בין אותה שיטה שמודדת תכונות שונות.

תוקף גורמי (פקטוריאלי) מבטא שתי תפיסות שונות לבדיקת תוקף המבנה: באמצעות ניתוח גורמים מגשש (exploratory factor analysis - EFA) ניתן לבדוק כמה גורמים (מבנים תיאורטיים) קיימים במדד ואת הקשר של פריטים בשאלון למבנה, או למבנים אלו; אם יש השערות מראש על מספר המבנים התיאורטיים במדד ועל שיוך של פריטים במדד לכל מבנה, ניתן לערוך ניתוח גורמים מאשש (confirmatory factor analysis - CFA).

תוקף תלוי קריטריון (criterion related validity)

תוקף תלוי קריטריון מבטא את הקשר בין מידע שהופק ממדד בשאלון ובין נתונים עובדתיים שניתן לסמוך שהם מהימנים ותקפים. ניתן למדוד תוקף תלוי קריטריון בשלוש דרכים: תוקף מקביל, תוקף ניבוי ותוקף בדיעבד. ניתן להשתמש בתוקף מקביל (concurrent validity) אם נתוני הקריטריון נאספים בעת איסוף המידע בסקר, למשל: השוואה בין דיווח על שכר של עובדים ובין נתוני הארגון על גובה השכר. תוקף ניבוי (predictive validity) משקף את היכולת של המדד בסקר לנבא נתוני קריטריון שייאספו לאחר עריכת הסקר, למשל: המתאם בין ציונים בבחינה פסיכומטרית ובין ציונים בסיום תואר ראשון. ככל שהמתאם בין נתוני המדד ובין נתוני הקריטריון גבוה יותר כך התוקף גבוה יותר. סוג תוקף נוסף, אך פחות נפוץ, הוא תוקף בדיעבד (postdictive validity), כאשר השימוש במדד בסקר נעשה לאחר שנתוני הקריטריון קיימים, למשל: מבחן פסיכולוגי שנערך לבחינת עוצמה של הלֵמֶת (הפרעת דחק פוסט-טראומטית).

מהימנות פנימית (internal reliability)

לאחר שנערך תיקוף תוכן לפריטים והוכח שהפריטים במדד רלוונטיים וממצים את התחום הנבדק, יש לבדוק את המהימנות הפנימית של המדד. מהימנות פנימית מבטאת את מידת העקביות הפנימית בין שאלות, שנועדו למדוד את אותו תחום מהיבטים שונים וכן את רגישות כלי המדידה לטעויות מקריות שמקורן בתוכן השאלות. למעשה, ככל שתוקף התוכן גבוה יותר (כלומר, פריטי המדד קשורים כולם לאותו התחום) מצופה כי המהימנות הפנימית תהיה גם היא גבוהה יותר (זאת בתנאי שהמושג או התכונה שנמדדים אינם רחבים מידי). מקדם המהימנות הנפוץ הוא מקדם אלפא של קרונברך (Cronbach's alpha).

במאמץ להגיע למהימנות פנימית טובה יש להקפיד שלא לפגוע בתוקף המבנה. (למשל: השמטת פריטים רלוונטיים לתחום התוכן, שקשורים בצורה בינונית למדד עשויה להעלות את ערך המקדם של המהימנות אך לפגום בכיסוי תחום התוכן.)

מהימנות מבחן חוזר (repeatability or test-retest reliability)

בדיקת מהימנות מסוג זה נותנת מידע על מידת היציבות של המדידה לאורך זמן ומשקפת את מידת הרגישות של כלי המדידה לטעויות מקריות שמקורן בזמן שחלף. בבדיקה נערכת השוואה בין המענה שנתן משיב לפריט בשאלון ובין המענה החוזר לאותו פריט לאחר זמן, באמצעות חשבון מקדם. ככל שהמקדם גבוה יותר כך כלי המדידה יציב יותר ורגיש פחות לשינויים בגורמים שקשורים בזמן (למשל: שינוי פנימי כגון שינוי במצב הרוח או שינוי בתנאים חיצוניים כגון במזג האוויר). הבדיקה מתאימה ל"תכונות" יציבות, שאינן משתנות בפרקי זמן קצרים. את הבדיקה החוזרת יש לערוך בפער זמן שאינו קצר מידי (כדי למנוע זכירת תשובות מבדיקה ראשונה) ואינו ארוך מידי (כדי שלא יהיו שינויים משמעותיים במאפיינים של המשיב); ומקובל לערוך את הבדיקה החוזרת לאחר כשלושה שבועות.

מקורות להטיה בפיתוח שאלון¹²

נוסף לדרישה שכלי המדידה יהיה תקף ומהימן, כל אחד ממרכיביו (פריטי השאלון) אמור להפיק מדידה חסרת הטיה. המדידה תהיה מדויקת או מוטית כתלות בדרך שבה השאלון יעוצב, בניסוח השאלות וההוראות, בדרך איסוף המידע ובתגובת המשיבים. איתור ההטיות נעשה באמצעות בדיקות שונות שנעשות לשאלון ולמערך האיסוף בשלבים שונים של פיתוח השאלון (כפי שיתואר בהמשך).

¹² להרחבה ראו:

Blair, J., Menon, G., & Bickart, B. (1991). Measurement effects in self vs. proxy response to survey questions: An information processing perspective. In P. B. Biemer, R. B. Groves, L. E. Lyberg, N. A. Mathiowetz, & S. Sudman (Eds.), *Measurement errors in surveys* (pp. 145–166). New York: Wiley.

Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied cognitive psychology*, 5(3), 213-236.

Schwarz, N., & Oyserman, D. (2001). Asking questions about behavior: Cognition, communication, and questionnaire construction. *The American Journal of Evaluation*, 22(2), 127-160

Tourangeau, R. (1984). Cognitive sciences and survey methods. In T.B. Jabine, M.L. Straf, J. M. Tanur, & R. Tourangeau (Eds.), *Cognitive Aspects of Survey Methodology: Building a Bridge between Disciplines* (pp. 73–101). Washington, DC: National Academy Press.

Tourangeau, R., & Bradburn, N. M. (2010). The psychology of survey response. In P. V. Marsden & J. D. Wright (Eds.), *Handbook of Survey Research* (pp. 315-346). Bingley: Emerald Group Publishing.

הטיות שמקורן בשאלון ובתהליך המענה

הבנת ההוראות, השאלות ואפשרויות התשובה, עשויות להיות מושפעת מההקשר שלהן בשאלון, למשל: סדר השאלות, טווח הזמנים המבוקש, מבנה התשובות ועוגני התשובה, דילוג על הוראות או קריאה חלקית שלהן, אורך השאלון וכדומה. הבנה נוגעת לשני היבטים: משמעות מילולית של השאלה והבנת כוונתה של השאלה.

הבנת המשמעות המילולית של השאלה קשורה לזיהוי המילים, לזיהוי של מושגים ולהיזכרות במידע ששמור בזיכרון הסמנטי של המשיב. בעיות בהבנת השאלה עשויות להתרחש כאשר המשיב לא מכיר מונחים או מושגים שרשומים בשאלה, כאשר השאלה מנוסחת בלשון גבוהה או עמומה, כשהיא כוללת יותר מנושא אחד, או כשהיא ארוכה או בעלת תחביר מורכב. בעיית הבנה עשויה להיות גם ביחס לזיהוי מוקד השאלה ולהבנת כוונתה. אם המשיב אינו בטוח במשמעותה או בכוונתה של השאלה, הוא עשוי לפרש את משמעותה או את כוונתה מתוך ההקשר (למשל: תכנים של שאלות סמוכות, אפשרויות התשובה, הסברים שנתן המראיין) ולהתמקד בנושאים שלהערכתו הם מרכזיים בשאלון, דבר שעלול להוביל לפרשנות שגויה של השאלה.

יש שאלות שדורשות איתור ושליפה של מידע מהזיכרון לטווח ארוך, לעתים תוך הסתייעות באסטרטגיות וברמזים. חלק מפריטי המידע שעליהם נשאלים בשאלונים מאוכסנים בזיכרון כמידע אירועי (אפיזודי) ייחודי, למשל: מאורעות, התנהגויות, תחושות ורגשות. חלק אחר מהמידע מאוכסן בזיכרון הסמנטי, למשל: ידע על העולם, מושגים, עובדות ושימוש בשפה. בזיכרון לטווח ארוך מאוכסנים פרטי מידע וידע עם קשרים הדוקים או רופפים, בהתאם לסוג הזיכרון ולדרך הקידוד והאכסון של המידע. חלק מתהליכי ההיזכרות הוא אוטומטי, אך חלק ניתן לתפעל בצורה יזומה, למשל באמצעות שימוש ברמזים לגבי מקומות ותקופות בזמן. בתהליך היזכרות קשה להבחין בין מאורעות דומים ויומיומיים וקל יותר להיזכר במאורעות נדירים. נוסף לגורמי הקשר שונים, שעלולים להפריע להיזכרות, עלול להפריע גם קישור של פריט המידע המבוקש לזיכרונות אחרים, שהינם קרובים בזמן או דומים במהות. בנוסף, משיב עלול לדווח בטעות על מאורעות שהתרחשו לפני התקופה המבוקשת (forward telescoping) או להשמיט אירועים שהתרחשו בטווח הזמנים המבוקש ולייחס אותם לתקופה מוקדמת יותר (backward telescoping).

כאשר ההיזכרות אינה מספקת תשובה שלמה ומדויקת, המשיב נדרש להפעיל תהליכי שיפוט, שכוללים השלמות למידע חסר או חלקי, שילוב של חלקי מידע לתשובה כוללת ושימוש בחישובים ואומדנים (למשל: אקסטרפולציה מתקופת זמן קרובה וקצרה לתקופת זמן רחוקה וארוכה). לעתים נעשה שימוש בהיוריסטיקות, מנגנונים שנועדו לקיצור תהליך ההשבה והפחתת הנטל הקוגניטיבי, כגון זמינות (נטייה לשפוט תדירות של אירועים בהתאם לקלות שבה מצליחים להיזכר בדוגמאות מסוימות או בכמות הדוגמאות); ייצוגיות (נטייה לשייך פריט מידע לקטגוריה מסוימת בהתאם לדמיונו לסטריאוטיפ של הקטגוריה); עיגון והתאמה (שימוש בשיפוט ראשוני ועריכת התאמה לא מיטבית,

שלרוב מוטה כלפי השיפוט הראשוני). השימוש בהיוריסטיקות מוביל לכך שמשיבים עלולים לתת אומדן בהתבסס על מדגמים קטנים ולא מייצגים או ממדדים לא תקפים, לנסות לאתר בזיכרון עדויות שיאששו דעה שנקבעה מראש ועוד.

במקרה שהמשיב נדרש למסור מידע על נחקר אחר (למשל: הורה על ילד, מטפל על מטופל וכו'), דיווח על עובדות ואירועים שהמשיב נחשף אליהם קל יותר לעומת מצבים שהמשיב לא היה עד להם ודיווח על עמדות ורגשות של האחר קשה יותר או בלתי אפשרי. אם המידע הנדרש חסר או חלקי, המשיב עשוי להסתייע באסטרטגיות נוספות לשם השלמת המידע. למשל: המשיב עשוי להשתמש לשם השוואה בעמדותיו, בתפיסותיו או בהתנהגויותיו שלו עצמו במידה שהוא תופס את עצמו כדומה לנחקר. המשיב עשוי להניח הנחות על טבעו של הנחקר, כך שהדיווח יתבסס על רושם כללי ועל תכונות שמיוחסות לנחקר במקום על התנהגויות או עמדות ספציפיות שנחקרות בסקר. בנוסף, אם הסקר בודק אירועים שהתרחשו בעבר, המשיב עשוי להניח כי תכונותיו והתנהגותו של הנחקר הן יציבות ולהתבסס על מצב נוכחי של הנחקר כאומדן למצב בעבר.

בסופו של התהליך הקוגניטיבי המשיב נדרש לבחור את התשובה לשאלה. בחירת התשובה מושפעת משלבים קודמים – הבנה, זכירה ושיפוט, אך גם ממאפיינים ייחודיים של השאלה עצמה, למשל: שאלה פתוחה או סגורה, שאלה על עובדות ואירועים או על עמדות ותחושות, מספר וסוג עוגני התשובה ועוד. שאלון ארוך ומורכב עלול להוביל את המשיב לשימוש ב"קיצורי דרך" ובסטיספיקציה (satisficing), למשל: בחירה בתשובה הראשונה שנראית מתאימה, הגיונית ו/או המהירה ביותר, בחירה בתשובה שאינה "מחייבת" (למשל: "לא יודע") או דילוג על השאלה. בנוסף, שאלות שעוסקות בנושאים רגישים (כלומר, שהתשובה עליהן סותרת נורמות חברתיות או משקפת עברה על החוק) עלולות להפיק תשובות מוטות עקב רצייה חברתית והרצון להצטייר באופן חיובי¹³.

¹³ להרחבה ראו:

Cerri, J., Davis, E. O., Veríssimo, D., & Glikman, J. A. (2021). Specialized questioning techniques and their use in conservation: A review of available tools, with a focus on methodological advances. *Biological Conservation*, 257, 1–19.

Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: a literature review. *Quality & quantity*, 47(4), 2025–2047.

Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological bulletin*, 133(5), 859.

יש סקרים שבהם המידע שנמסר על נדגם אינו מתקבל באופן ישיר ממנו, אלא באמצעות מרואיין חליפי (proxy). מרואיין חליפי עשוי להיות נציג של משק הבית שעונה עבור בני המשפחה, מטפל שעונה עבור מטופל (בבית או במוסד), הורה על קטין ועוד. דיווח באמצעות מרואיין חליפי נפוץ במפקדי אוכלוסין, בסקרי כוח אדם, בסקרי הוצאות משקי בית ובסקרי בריאות. לשימוש במרואיין חליפי יש יתרון בהעלאת שיעור השבה, בהפחתת נטל ההשבה ובחיסכון בעלויות איסוף המידע. לצד שיקולי נוחות וחיסכון, השימוש במרואיין חליפי נעשה לעיתים מתוך הכרח, כאשר הנדגם אינו יכול למסור מידע, למשל: במקרה של נדגמים קטינים, נדגמים עם מוגבלות שאינה מאפשרת השתתפות בסקר וכדומה. מרואיין חליפי עשוי להוות תחליף מהימן לנדגם כאשר הסקר

¹⁴ להרחבה ראו:

Abate, C., Lippé, S., Bertout, L., Drouin, S., Krajcinovic, M., Rondeau, É., ... Sultan, S. (2018). Could we use parent report as a valid proxy of child report on anxiety, depression, and distress? A systematic investigation of father–mother–child triads in children successfully treated for leukemia. *Pediatric blood & cancer*, 65(2).

Cobb C. (2018). Answering for someone else: Proxy reports in survey research. In D. Vannette, J. Krosnick (Eds.), *The Palgrave Handbook of Survey Research* (pp. 87-93). Palgrave Macmillan, Cham.

Khadka, J., Kwon, J., Petrou, S., Lancsar, E., & Ratcliffe, J. (2019). Mind the (inter-rater) gap. An investigation of self-reported versus proxy-reported assessments in the derivation of childhood utility values for economic evaluation: A systematic review. *Social Science & Medicine*, 240, 112543.

McPherson, C. J., & Addington-Hall, J. M. (2003). Judging the quality of care at the end of life: Can proxies provide reliable information?. *Social science & medicine*, 56(1), 95-109.

Schwarz, N., & Oyserman, D. (2001). Asking questions about behavior: Cognition, communication, and questionnaire construction. *The American Journal of Evaluation*, 22(2), 127-160.

Schwarz, N., & Wellens, T. (1997). Cognitive dynamics of proxy responding: The diverging perspectives of actors and observers. *Journal of Official Statistics*, 13(2), 159-179.

בודק עובדות שגלויות וידועות הן לנדגם והן למרואיין החליפי. אולם במקרים רבים עשוי להיות פער בין שני מקורות מידע אלה, דבר שעלול לגרום להטיות באומדנים.

בהתאם למודל ההשבה הקוגניטיבי, לרוב לא צפוי הבדל בהבנה כשמדובר במרואיין חליפי, אולם הוא עשוי שלא להכיר מושגים ומונחים שידועים לנדגם (למשל: בגלל פערי גיל, אי-הכרת מושגים קשורים לעבודתו של הנדגם ועוד). מרואיין חליפי שלא נחשף באופן ישיר לאירועים שקשורים לנדגם יתקשה לדווח עליהם באופן מהימן; אם המידע נרכש דרך שיחה הוא יאוכסן כזיכרון סמנטי כללי והדיווח יושלם בהערכה (estimation) ולא הייתה כל חשיפה למידע הוא לא יקודד בזיכרון ולא ניתן יהיה לדווח עליו. ככל שהמרואיין החליפי נמצא יותר זמן עם הנדגם כך הוא חשוף יותר להתנהגויות ולמצבים שונים שלו וכך הדיווח שלו יהיה דומה יותר לדיווח עצמי, אך לא כשמדובר ברגשות ובתחושות פנימיים (למשל: כאב, דיכאון, חרדה) של הנדגם.

מרואיין חליפי ייטה להשתמש בהנחות לגבי יציבות ולכן הוא ייטה להסתמך יותר על רושם כללי ותכונות של הנדגם מאשר על התנהגויות ספציפיות שלו; הוא יטה גם לאמוד את מצבו והתנהגותו בעבר על פי מצב נוכחי והתנהגות עכשווית של הנדגם. השימוש בהנחות לגבי יציבות מסביר את העקביות הגבוהה שקיימת בתשובות של מרואיינים חליפיים על שאלות שעוסקות בנושאים דומים אצל נדגם ("מהימנות פנימית") ולאורך זמן ("מהימנות חוזרת"), אך הדבר אינו מעיד בהכרח על תוקף התשובות.

בשאלות על נושאים עם רגישות חברתית נדגם ייטה, יותר ממרואיין חליפי, להצגה עצמית בהתאם לנורמות חברתיות. עם זאת, רצייה חברתית עלולה להתערב בבחירת התשובה גם אצל מרואיין חליפי, למשל: כשהורה נשאל על צריכת אלכוהול של ילדו, על אלימות בבית כלפי הילד וכדומה.

לבסוף, אם מרואיין חליפי נדרש לדווח על מספר אנשים ממשק הבית, עליו לענות על מספר רב של שאלות (שחלקן חוזר על עצמו). נוסף לזמן הנוסף שיהיה עליו להקדיש למענה, הוא יידרש במקרים רבים להשתמש בהערכות ואומדנים. מצבים כאלה דורשים מאמץ קוגניטיבי רב, דבר שמגביר את הסיכון לשימוש ב"קיצורי דרך" (satisficing).

מקורות להטיה באיסוף הנתונים

שיטת איסוף הנתונים¹⁵

איסוף הנתונים עשוי להתבצע בשיטות (modes) שונות, למשל: סקר שנערך פנים אל פנים או באמצעות הטלפון על ידי מראיינים (סוקרים), באמצעות שאלוני נייר או שאלונים מקוונים; שאלון שנשלח בדואר רגיל, בדואר אלקטרוני, דרך אתר באינטרנט או בפנל אינטרנטי; או שילוב של כל אלה (mixed-modes) בסקר למשיבים שונים או לאותו משיב, למשל בפניות שונות או בגלים שונים. בחירה בשיטת איסוף מסוימת, או בשילוב של מספר שיטות, צריכה להיעשות בהתחשב בהיבטים שונים של הסקר (אורך השאלון ומורכבותו, רגישות התכנים, מאפייני אוכלוסיית הסקר והמשיבים, גודל התקציב) וביתרונות ובחסרונות של כל שיטה, למשל: השבה עצמית על שאלון שמכיל נושאים רגישים תפיק תשובות פחות ופחות רצייה חברתית לעומת השבה באמצעות מראיין. לעומת זאת, בריאיון פנים אל פנים ניתן להשיב על שאלון ארוך, יותר מאשר בהשבה עצמית. (ראו לוח 1.1)

ככלל, יש להימנע משימוש בדרכים לאיסוף מידע מאוכלוסייה שאין ברשותה אמצעים או מיומנויות לשימוש בהן (למשל: מענה מקוון מחייב נגישות למחשב או לטלפון נייד ומיומנות בשימוש בהם). לעתים נעשה שילוב של שיטות (mixed-modes)¹⁶ במטרה להעלות את שיעור ההשבה לסקר. אם נדגמים לא משיבים לפנייה בדואר אלקטרוני הם יקבלו פנייה בטלפון להשתתף בסקר ולעתים אף פנייה לראיון בבית. איסוף הנתונים בדרך של פנים אל פנים יקר יותר, אך מעלה באופן משמעותי את שיעורי ההשבה ואף מאפשר ייצוג טוב יותר של קבוצות מסוימות באוכלוסייה.

¹⁵ להרחבה ראו:

Cernat, A., & Sakshaug, J. (2020). The impact of mixed modes on multiple types of measurement error. *Survey Research Methods*, 14(1), 79-91.

Daikeler, J., Bosnjak, M., & Lozar Manfreda, K. (2020). Web versus other survey modes: An updated and extended meta-analysis comparing response rates. *Journal of Survey Statistics and Methodology*, 8(3), 513-539.

de Leeuw, E. D. (2018). Mixed-mode: Past, present, and future. *Survey Research Methods*, 12(2), 75-89.

Schork, J., Riillo, C. A., & Neumayr, J. (2021). Survey mode effects on objective and subjective questions: Evidence from the Labour Force Survey. *Journal of Official Statistics*, 37(1), 213-237.

¹⁶ להרחבה ראו:

Sakshaug, J. W., T. Yan, and R. Tourangeau (2010). Nonresponse error, measurement error, and mode of data collection tradeoffs in a multi-mode survey of sensitive and non-sensitive items. *Public Opinion Quarterly*, 74, 907-933.

לוח 1. יתרונות וחסרונות של דרכים שונות לאיסוף נתונים

סוג שאלון	יתרונות	מגבלות
שאלון נייר – השבה עצמית בדואר רגיל	• אין לחץ זמן בהשבה; • אפשר לכלול הוראות מפורטות, אפשרויות תשובה רבות וחומרי עזר (כגון מפות ותמונות); • פחות רצייה חברתית ותשובות אמינות יותר לשאלות רגישות.	• למשיב צריך להיות קל להבין ולהתמצא (למשל: משפטים קצרים וברורים; עיצוב מרווח; ללא מעברים מורכבים); • השאלון צריך להיות קצר, יחסית; • קושי להשיב על שאלות במבנה לוח; • המשיב נדרש לכישורים טובים של קריאה וכתיבה (בשאלות פתוחות); • אין בקרה על איכות המידע בזמן השבה.
שאלון נייר – ריאיון פנים אל פנים	• ניתן לערוך ריאיון ארוך; • ניתן לשאול שאלות מורכבות ולהשתמש בחומרי עזר.	• נדרשת הכשרה מעמיקה לסוקר, כיוון שהוא האמצעי היחיד לבקרה ולמניעה של טעויות השבה; • עלולה להיות הטיה בהשבה על שאלות רגישות ו/או שמעוררות רצייה חברתית.
שאלון ממוחשב [כללי]	• ניתן להשתמש בשאלות ניתוב ולשלב בקורות למניעת טעויות וחוסר עקביות בתשובות; • ניתן לשלב השלמה אוטומטית לתשובות לשאלות פתוחות פשוטות.	• למשיב צריך להיות קל לנווט בשאלון; • רצויה אפשרות לקבלת סיוע מקוון; • עבודת התכנות ובחינת המערכת עשויות להיות מורכבות, ארוכות בזמן ויקרות.
שאלון ממוחשב – השבה עצמית [נוסף לעיל]	• ניתן לשלב חומרי עזר.	• למשיב צריך להיות קל להבין ולהתמצא; • השאלון צריך להיות קצר, יחסית; • קושי להשיב על שאלות במבנה לוח.
שאלון ממוחשב – ריאיון טלפוני [נוסף לעיל]	• ניתן לשלב מידע ממקור אחר.	• נדרשות שאלות קצרות ומעט אפשרויות תשובה; • קושי לשלב שאלות פתוחות; • יותר רצייה חברתית לעומת השבה עצמית.
שאלון ממוחשב – ריאיון פנים אל פנים [נוסף לעיל]	• ניתן לערוך ריאיון די ארוך; • ניתן לשלב הוראות והגדרות מפורטות וחומרי עזר.	• יותר רצייה חברתית לעומת השבה בטלפון.
מקור: מותאם מברנקטו ועמיתיה (Brancato et al., 2006)		

לעתים נעשה מעבר בין שיטות כדי להגביר את דיוק המדידה. אחת הדוגמאות לכך היא במצב שבו השאלון עוסק בנושאים רגישים: תחילה נעשית פנייה לנדגם בטלפון על מנת להשיג את הסכמתו להשתתף בסקר ולקבל פרטי רקע נחוצים ולאחר מכן הוא מתבקש לענות על שאלון הסקר שיישלח

אליו בדואר האלקטרוני או באמצעות קישור באינטרנט, כדי שהתשובות יהיו כנות יותר ומדויקות יותר. המחקר מראה¹⁶ כי מעבר בין ריאיון ובין השבה עצמית מעלה את דיוק התשובות, אך יש מקרים שהוא מוביל לנשירה (אי-השבה ברמת היחידה).

שימוש במראיינים¹⁷

יש להקדיש תשומת לב להטיות אפשריות שנובעות משימוש במראיינים (סוקרים) לצורך ביצוע הסקר. מראיין תורם להעלאת שיעור ההשבה אך עלול לתרום להטיות ייצוג, להטיות מדידה (ואף להטיות עיבוד אם המראיין נדרש להקליד קודים מספריים בעת השבה על שאלות פתוחות). מחקרים מראים כי מראיינים מנוסים ומראיינים עם גישה חיובית (למשל: אופטימיסטים) מקבלים שיתוף פעולה רב יותר בסקרים. מראיין מנוסה יודע להתאים את הפנייה למראיין, לשמור על אינטראקציה בשיחה ולשכנע – דבר שמוביל לשיתוף פעולה. יש לשים לב שיכולות אלה לא ייצרו הטיית-השבה לא רצויה, כלומר קבלת שיתוף פעולה גבוה מאוד רק בקרב תת-קבוצה מובחנת באוכלוסייה (למשל: בקרב קשישים). נוסף לכך, יש להיזהר ממצב שמראיין עלול לתרום להטיית כיסוי אם הוא יטה לסווג נדגמים "קשים" כ"מקרי אפס" במקום להשקיע מאמץ ולבדוק לעומק אפשרות שהנדגמים שייכים לאוכלוסייה.

נוסף לאלה, מראיין עלול להגביר הטיות מדידה. ראשית, יש שאלות שנוטות יותר לגרור השפעה של המראיין על המדידה, למשל: שאלות על עמדות, שאלות על נושאים רגישים, שאלות מורכבות ושאלות פתוחות, שלעיתים דורשות תיווך (הסבר או פרשנות) של המראיין. מצב זה עלול להוביל לכך שהמראיין יקריא הוראות ושאלות באופן שונה למשיבים שונים. בנוסף, לעיתים חוסר התאמה בין מאפייני הנדגם למאפייני המראיין (כגון שפה, רקע תרבותי) עלול להוביל לחוסר הבנה של רמזי התנהגות ושפה ובכך לתרום להטיות מדידה. בסך הכול, יש משתנים רבים שעשויים לתרום להטיות, למשל: מאפייני רקע של המראיין; ניסיון והכשרה של המראיין; גישה של עורכי הסקר (כגון תגמול בהתאם לשיעורי השבה); עומס שמוטל על המראיין; עמדות, אישיות וכישורים חברתיים של המראיין; מאפיינים של המראיין ומאפייני השאלון.

¹⁷ להרחבה ראו:

Daikeler, J., & Bosnjak, M. (2020). How to conduct effective interviewer training: A meta-analysis and systematic review. In: K. Olson, J. D. Smyth, J. Dykema, A. L. Holbrook, F. Kreuter & B. T. West (Eds). *Interviewer Effects from a Total Survey Error Perspective*. Boca Raton: CRC Press, pp. 47– 62.

Mittereder, F., Durow, J., West, B. T., Kreuter, F., & Conrad, F. G. (2018). Interviewer–respondent interactions in conversational and standardized interviewing. *Field Methods*, 30(1), 3-21.

West, B. T., & Blom, A. G. (2017). Explaining interviewer effects: A research synthesis. *Journal of survey statistics and methodology*, 5(2), 175-211.

חלק מהמאמצים למנוע הטיות בסקר נעשה באמצעות הכשרה למראיינים. מומלץ להכשיר מראיינים למנוע סירובים באמצעות אינטראקציה נכונה עם נדגמים (יותר מאשר לשכנע סרבנים), להיצמד להנחיות ולהוראות מובנות במטרה להפחית הטיות מדידה ולכלול בהכשרה תרגול ומתן משוב.

מתן תגמול לעידוד ההשבה¹⁸

הצעת תגמול (תמריץ) קטן למשיבים הוא הליך מקובל שננקט כדי להעלות את שיעורי ההשבה בסקרים. מחקרים מראים באופן עקבי כי מתן תגמול הוא הליך יעיל להעלאת שיעורי ההשבה, בעיקר תגמול שניתן מראש (לעומת תגמול שמותנה במענה לסקר) וכי תשלום כספי יעיל יותר מתגמולים אחרים (כגון שוברי קנייה ומתנות). ממצא זה חוזר על עצמו בסקרים ובניסויים שבוצעו בשיטות איסוף שונות. עוד נמצא כי העלאת שיעור ההשבה משמעותי יותר בסקרי דואר ובסקרי טלפון ופחות משמעותי בסקרי פנים אל פנים (שבהם שיעורי ההשבה ממילא גבוהים יותר); מתן תגמול מוסיף לשיעורי ההשבה גם בפנלים הסתברותיים ובפנלים לא-הסתברותיים שבהם שיעורי ההשבה גבוהים יחסית.

מחקרים שנעשו על מתן תגמול בשיטות איסוף שונות לא מצאו כי תגמול תרם לעליה (או לירידה) בהטיית אי-השבה ולא נמצא כי הוא מגביר הטיות מדידה. עם זאת, אין די מחקר בנושא ולכן יש להשתמש במתן תגמול באופן מבוקר ותוך בדיקת השפעות אפשריות שלו על הטיות ייצוג ומדידה.

¹⁸ להרחבה ראו:

Mercer, A., Caporaso, A., Cantor, D., & Townsend, R. (2015). How much gets you how much? Monetary incentives and response rates in household surveys. *Public Opinion Quarterly*, 79(1), 105-129.

Oscarsson, H., & Arkhede, S. (2020). Effects of conditional incentives on response rate, non-response bias and measurement error in a high response-rate context. *International Journal of Public Opinion Research*, 32(2), 354–368.

Schwarz, H., Revilla, M., & Struminskaya, B. (2022). Do previous survey experience and participating due to an incentive affect response quality? Evidence from the CRONOS panel. *Journal of the Royal Statistical Society Series A*, 1-23.

Shamon, H., & Berning, C. (2020). Attention check items and instructions in online surveys with incentivized and non-incentivized samples: Boon or bane for data quality?. *Survey Research Methods*, 14(1), 55-77.

Stanley, M., Roycroft, J., Amaya, A., Dever, J. A., & Srivastav, A. (2020). The effectiveness of incentives on completion rates, data quality, and nonresponse bias in a probability-based internet panel survey. *Field Methods*, 32(2), 159-179.

התמודדות עם הטיות במדידה ובאיסוף

חשוב לאתר מראש גורמים שעשויים לתרום להטיות ולהתמודד איתם, בטרם נערך הסקר. מדידה לא מדויקת ומוטה תגדיל את הסטייה המקרית של המדידה, דבר שעלול להוביל לפגיעה ביכולת להסיק מסקנות תקפות מהסקר, כגון אומד של פרמטר באוכלוסייה או קיום של קשר בין משתנים. להלן ייסקרו שיטות שונות לזיהוי ולבדיקה של הטיות בשלבים שונים של פיתוח הסקר – בדיקות מקדימות, בדיקות בשדה ובדיקות לאחר איסוף הנתונים¹⁹. כפי שהודגש בתחילת הנספח, המטרה היא לתת תמונה על מגוון שיטות מרכזיות, כאשר על מזמין הסקר ו/או מבצעו מוטל להחליט אלו מבין השיטות ראויים להיבחר לסקר. ככלל, מומלץ לשלב שיטות בדיקה של שלבים שונים על מנת להגביר את היעילות בהתמודדות עם ההטיות האפשריות.

בדיקות מקדימות

בדיקות מקדימות נעשות טרם איסוף הנתונים בשדה, במטרה לאתר הטיות במדידה, הן איכותניות באופיין ואין דורשות שימוש במדגמים גדולים. למרות שהבדיקות מתאימות בעיקר לשלבים הראשונים של פיתוח השאלון, ניתן להשתמש בהן גם בשלבים מתקדמים יותר ורצוי להשתמש ביותר מסוג אחד של בדיקות מקדימות. שתי בדיקות מקדימות מרכזיות הן סקירה מתודולוגית של מומחים וראיונות קוגניטיביים. בדיקות נוספות כוללות שימוש בקבוצות מיקוד, בדיקה לא-פורמאלית של עמיתים, בדיקות היתכנות ובדיקות שימושיות לשאלונים ממוחשבים.

¹⁹ להרחבה ראו:

Beatty, P. C., Collins, D., Kaye, L., Padilla, J. L., Willis, G. B., & Wilmot, A. (Eds.). (2019). *Advances in questionnaire design, development, evaluation and testing*. John Wiley & Sons, Incorporated.

Brancato, G., Macchia, S., Murgia, M., Signore, M., Simeoni, G., Blanke, K., & Hoffmeyer-Zlotnik, J. (2006). *Handbook of recommended practices for questionnaire development and testing in the European statistical system*. European statistical system.

McCarthy, J. S., Ott, K., Ridolfo, H., McGovern, P., Sirkis, R., & Moore, D. (2018). Combining multiple methods in establishment questionnaire testing: The 2017 census of agriculture testing bento box. *Journal of Official Statistics*, 34(2), 341-364.

Forsyth, B. H. & Lessler, J. T. (1991). Cognitive laboratory methods: A taxonomy. In P. P. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz, & S. Sudman (Eds.), *Measurement errors in surveys* (pp. 393–418). New York: Wiley-Interscience.

Jobe, J. B. (2003). Cognitive psychology and self-reports: Models and methods. *Quality of Life Research*, 12(3), 219-227.

Willis, G. B. (2015). *Analysis of the cognitive interview in questionnaire design*. Oxford University Press.

סקירה מתודולוגית של השאלון נעשית במטרה להעריך את הקשיים האפשריים בשאלון למשיב ולמראיין. הסקירה נעשית על ידי חוקרים בעלי מומחיות בתכנון ובפיתוח שאלונים. לרוב, הסקירה נעשית בשלבים מוקדמים של פיתוח השאלון, לעתים בשילוב עם בדיקות מקדימות אחרות. בסקירה המומחה בודק בעיות אפשריות שקשורות לתהליכי המענה לשאלון.

סקירה מתודולוגית של מומחים יכולה להיעשות באמצעות מערכת קידוד להערכה קוגניטיבית של השאלות והיא יכולה להוות חלק מבדיקת שימושיות של שאלון ממוחשב. לחילופין, הסקירה יכולה לכלול ניתוח קוגניטיבי לא מובנה של המענה לשאלון, כולל הערכת השינויים הנדרשים בשאלון לשם מענה מיטבי על השאלות.

שימוש בראיונות קוגניטיביים נעשה על מנת לאתר קשיים בתהליך הקוגניטיבי בעת המענה לשאלון ולהסתייע בתובנות של המשיבים לגבי פתרונות אפשריים. המשיבים עשויים להתבקש לתאר בקול רם את מחשבותיהם בזמן מענה על השאלון או בסיום המענה. ניתן לשלב בריאיון שאלות חקירה כדי להעמיק בנושאים ובהיבטים בשאלון שעלולים להיות בעייתיים או כאלה שהתגלו בהם קשיים תוך כדי ריאיון. בנוסף, ניתן לבקש מהמשיב לסווג ולסדר נושאים שונים שהשאלון עוסק בהם. רכיב נוסף שניתן לכלול בראיונות הקוגניטיביים הוא חידוד הגדרות אופרציונליות למושגים באמצעות קטע קצר (vignette) שכולל מושג עם דוגמאות, והמשיב מתבקש לציין באיזו מידה כל דוגמה מבטאת וממצה את המושג. לבסוף, בריאיון הקוגניטיבי ניתן לבדוק את הזמן שלוקח למשיב לענות על שאלה (באמצעות מדידה עם שעון עצר או באמצעות הערכה איכותנית). אולם, כיוון שזמן התגובה הוא אינדיבידואלי, יש ליצור "מדד בסיס", באמצעות מדידת זמן התגובה לשאלות ניטרליות.

קבוצות מיקוד עשויות לסייע בשלבים הראשונים של פיתוח השאלון, במיפוי הידע על נושאי החקירה ובגיבוש רעיונות לשאלות. בהמשך, ניתן להיעזר בקבוצות מיקוד לשם הבנה לגבי תהליך ההשבה על השאלון, הבנת מושגים ומונחים, תחושת קלות או קושי במענה וכן הלאה. קבוצות מיקוד עשויות לתת תובנות גם לגבי שיטות איסוף מתאימות והדרכות למראיינים במקרה שהסקר נעשה על קבוצות ייחודיות באוכלוסייה ו/או כשהשאלון עוסק בנושאים רגישים.

לאחר ניסוח טיוטה ראשונה של השאלון ניתן להיעזר בעמיתים לשם קבלת משוב ראשוני, בעיקר לגבי עיצוב השאלון, סדר השאלות, ניסוח השאלות, מעבר בין שאלות ותזרים השאלות, בעיות הגהה וכדומה. רצוי לקבל משוב מעמיתים שעוסקים במגוון תחומים ואף ממשיבים פוטנציאליים.

בדיקות גישוש (exploratory) או היתכנות (feasibility) נעשות לרוב באמצעות ביקורים באתר שבו ייערך הסקר (למשל: ארגון) והן משמשות לבדיקת ההתאמה בין הנושא שנבדק ובין הביטוי האופרטיבי שלו במציאות. בהקשר של מחקרים כלכליים, הבדיקות מתמקדות בעיקר בנתונים המבוקשים בסקר, למשל: זמינות הנתונים הרצויים ברשומות, זמני הרישום שלהם והגדרת המושגים כפי שהם מצויים ברשומות החברה. בנוסף, בדיקות אלה עשויות לסייע בהערכה של נטל ההשבה ובזיהוי של בעלי התפקיד המתאימים למענה לסקר.

לשאלונים מקוונים נעשות בדיקות נוספות, כאשר המטרה העיקרית שלהן היא לגלות ולסלק חסמים שמפריעים למראיינים ולמשיבים להשתמש בשאלון בצורה מדויקת ויעילה. בבדיקה יש להתמקד בבעיות שעלולות להוביל לדיווח שגוי, לאי-השבה ולהפסקת מענה, תוך סקירה של מגוון היבטים לרבות גופני הכתב, סמלים (icons), פריסה על הדף, הזנת נתונים וניווט. ראשית, השאלון נבדק מבחינת תפקוד ונראות (למשל: איתור עיצוב ויזואלי לא מיטבי, שימוש לא אחיד בתבניות ובגופנים, טעויות דקדוקיות, חוסר התאמה בין אפשרויות התשובה ובין הנדרש בשאלה, מעבר שגוי בין שאלות, הופעה שגויה של ההוראות לסוקר, ועוד). נוסף לכך, נערכות לשאלון בדיקות שמישות (usability testing), אשר מתמקדות בדרך שבה המראיין ו/או המשיבים משתמשים במערכת, למשל: אסטרטגיות כלליות שננקטות לשם השלמת המענה לשאלון; מעברים שונים שנעשים בין שאלות, אם יש שאלות סינון וניתוב; מקומות בשאלון שהמשיבים חשים בלבול או תסכול; טעויות שנעשות במענה, ניסיונות תיקון של המענה על ידי המשיב והצלחה או כישלון בתיקון המענה; צורך של משיבים בקבלת עזרה; איתור מצבים שלא יהיה ניתן להשלים את המענה לשאלון. את בדיקות השמישות ניתן לערוך בדומה לבדיקות שנעשות במסגרת ראיון קוגניטיבי למשתמשים (למשל: "חשיבה בקול רם" ושימוש בשאלות חקירה). בשלבים מוקדמים של תכנון השאלון, ניתן לערוך בדיקות שמישות באמצעות גרסה לא מתקדמת של אב טיפוס של השאלון (למשל באמצעות שימוש במסכי נייר) וככל שמתקדמים בפיתוח השאלון, ניתן לבחון היבטים נוספים של שימוש בשאלון כגון תכונות ניווט, אסטרטגיות של תיקון שגיאות, וכדומה. היבט נוסף בבדיקת השמישות של השאלון היא מידת התאמתו לאנשים עם מוגבלות.

בדיקות בשדה

מטרתן של בדיקות אלה לבדוק את הליך איסוף המידע בתנאי שדה. בדיקות בשדה חשובות במיוחד כשמדובר בסקרים מתמשכים או חוזרים. ניתן לערוך סקר מקדים (pretest), שכולל הדמיה של שימוש בכלי המדידה ושל הפעולות שכרוכות במערכת האיסוף, ניסוי לבדיקת השערות על הבדלים בניסוח פריטים ו/או בשיטות איסוף באמצעות ניסוי מבוקר, ניתוח של אינטראקציות מראיין-משיב ותחקור של מראיינים, משיבים וגורמים אחרים שמעורבים באיסוף הנתונים.

אחת הבדיקות המרכזיות בשדה היא עריכת סקר מקדים, שמאפשר לערוך "חזרה כללית" לסקר. בבדיקה זו מדמים את הסקר האמיתי, כאשר לבד מגודל המדגם, שאר מרכיבי הסקר זהים, לרבות תפעול בהתאם ללוח זמנים מתוכנן, שליחת הודעות מקדימות לסקר ותזכורות למענה, מעקב אחר התקדמות הסקר וטיפול במצבים של קושי באיתור נדגמים, סירוב להשתתף וכדומה. הסקר המקדים מאפשר לבדוק היבטים שקשה לאתרם בבדיקות מקדימות, למשל: איתור פריטים עם שיעור אי-השבה גבוה או עם התפלגות בעלת דפוס חריג ו/או שונות נמוכה, אופי הקשר בין פריטי מדד מסוים, נקודות חוזק וחולשה של שיטות איסוף חדשות ו/או של שילוב של שיטות שונות, איתור קשיים של מראיינים ובעיות באינטראקציה בין מראיינים לנדגמים, קושי בהודעות מקדימות על הסקר, בעיות בתפעול ובמענה לסקר מקוון ועוד. חיסרון מרכזי של הסקר המקדים הוא שלרוב הוא יקר יותר וגוזל

יותר זמן מהבדיקות המקדימות. בנוסף, אם הוא נעשה ללא שילוב עם בדיקות מקדימות, הוא עלול שלא לגלות בעיות של המשיב עם השאלון (למשל: הבנה, רגישות של שאלות, קושי לדלות מהזיכרון, קושי בבחירת תשובה, מוטיבציה נמוכה לענות).

שימוש בניסוי מבוקר בשדה (randomized experiment או split ballot) מאפשר להשוות בין גרסאות שונות של שאלונים (למשל: אורך וקצר), בין גרסאות שונות של פריטים בשאלון ובין שיטות איסוף (modes) שונות. על מנת להכריע איזו חלופה עדיפה רצוי תיקוף באמצעות נתונים אובייקטיביים. למשל: השוואה בין תשובות לשני נוסחי שאלה על מספר שעות עבודה בתקופת זמן מסוימת ובין נתונים שקיימים במערכת של הארגון על תקופת זמן זו). חשוב שהניסוי ייערך באמצעות הקצאה מקרית של המשתתפים לקבוצות השונות ועם מדגם גדול דיו, כך שניתן יהיה לערוך את כל ההשוואות הנדרשות.

ניתן לבצע קידוד אינטראקציה בין משיב ומראיין מתוך ראיונות חיים או מוקלטים, פנים אל פנים או בטלפון. מטרת הבדיקה דומה לבדיקות המקדימות – איתור בעיות בשאלון באמצעות התמקדות בהתנהגויות של המראיין ושל המשיב. השיטה מאפשרת לערוך זיהוי שיטתי של שאלות שכוללות מספר גדול של בעיות, אך לרוב היא אינה מאפשרת לזהות את מקור הבעיה ואינה מזהה בעיות שאין להן ביטוי התנהגותי כגון פרשנות לא נכונה של משמעות השאלה.

לאחר איסוף הנתונים ניתן להשתמש בשאלון מובנה על מנת לקבל מהמשיב מידע על פרשנותו למשמעות של שאלות הסקר. למשל: ניתן לבדוק אם משיבים הבינו באופן דומה שאלות ומושגים שונים בסקר ואם קיים פער בין כוונת מתכנן הסקר ובין פרשנות המשיבים, אם ישנן שאלות מיותרות בשאלון, אם נדרשות הבהרות למושגים ולמונחים ועוד. ניתן לערוך את התחקור בסיום המענה על הראיון, או באמצעות יצירת קשר עם המשיב לאחר שהסקר הסתיים. בנוסף, התחקור עשוי להצביע על מקור של בעיות, אם היו כאלה בסקר.

תחקור המראיינים (הסוקרים) מסייע במקרים שבהם תצפיות במראיינים ובמראיינים או תחקור של מראיינים לא שיקפו באופן מספק בעיות בהבנה של השאלון או בהעברה שלו. המטרה היא להיעזר בקשר הישיר שיש למראיינים עם המשיבים, על מנת להעשיר את ההבנה של מפתח הסקר לגבי בעיות בשאלון ולקבל תובנות על פתרונות אפשריים. התחקור יכול להיערך באמצעות שיחה קבוצתית עם המראיינים או באמצעות העברת שאלונים מובנים למראיינים.

בדיקות לאחר עריכת הסקר

לאחר סיום הסקר ניתן לקבל מידע נוסף על איכותו של השאלון ועל מערך האיסוף. הבדיקות נסמכות על ניתוחים סטטיסטיים, כגון: שיעורי אי-השבה לפריט, דפוסי התפלגות התשובות, שיעורי זקיפה ועוד. בנוסף, ניתן לכלול בדיקות איכותניות (למשל: באמצעות תחקור משיבים).

ניתן לערוך ראיונות חוזרים, זהים לראיונות שנעשו בסקר, למדגם של משיבים. מטרתם של הראיונות החוזרים היא להעריך את טעות המדידה, שעשויה לנבוע מהמשיב (למשל: חוסר הבנה, בלבול ועוד),

מהשאלון (למשל: ניסוח לא ברור של מושגים) ו/או מהמראיין. לאחר ביצוע הריאיון החוזר ניתן לערוך ניתוח לעקביות של התשובות ולאתר מענים שקיימים בהם הבדלים. במידת האפשר, ניתן להעמיק את החקירה ולבדוק מה הסיבות למענה שונה בריאיון החוזר.

לאחר הסקר יש לחקור לעומק את הסיבות לשיעורי אי-השבה גבוהים ואף לשיעורים נמוכים יותר כאשר מדובר בשאלות ליבה בסקר. סיבות אפשריות קשורות, למשל, לבעיה בתוכן (למשל: עיסוק בנושאים רגישים), במבנה השאלה ו/או אפשרויות התשובה, במיקום של השאלות בשאלון (למשל: בסוף השאלון), בשיטת האיסוף (mode), במשיב (למשל: המשיב הוא מראיין חליפי) ועוד.

נוסף למקרים של אי-השבה ניתן לבחון את דפוסי ההשבה לשאלות מסוימות (למשל: שונות נמוכה, ממוצע נמוך מאוד או גבוה מאוד), במקרים שיש חשש שההשבה עלולה להיות מוטה, במקרים שנעשה שינוי בניסוח השאלה ו/או באפשרויות התשובה ועוד.

שיעורים גבוהים של עריכת נתונים (data editing) ושל זקיפות (imputations) עשויים להעיד על בעיות מסוגים שונים בחלקים שונים של השאלון (למשל: ניסוח של שאלות, תזרים השאלות, הבנה של מושגים ומונחים בשאלון, הכשרת המראיינים ועוד).

המלצות לעיצוב השאלון ולניסוח שאלות הסקר²⁰

בשלב ראשון בפיתוח השאלון, על יוזמי הסקר להחליט אלו נושאים ומושגים יכללו בשאלון, לרוב לאור סקירת ספרות שנעשתה בנושא ובשילוב עם מומחי תוכן ומומחים מתודולוגיים. לאחר שהוסכם על הנושאים ועל המושגים שיכללו בסקר, הם "מתורגמים" לשאלות בשאלון. לרוב נעשה שימוש בשאלונים, סולמות או פריטים שפותחו בעבר, דבר שחוסך זמן ומשאבים, אך גם תורם לאחידות ומאפשר השוואות בין סקרים ולאורך זמן. יש לעשות מאמץ ולהשתמש בהגדרות מושגיות קיימות (למשל: הגדרות הלמ"ס למושגים ולמונחים) ובהגדרות ואופרציונליות קיימות (למשל: שאלונים בשימוש בסקרים בינלאומיים). נוסף לחיסכון בעלויות, שימוש במושגים, במונחים ובהגדרות משותפים תורם למהימנות ולתוקף של שאלון הסקר.

20 להרחבה ראו:

Dykema, J., Schaeffer, N. C., Garbarski, D., & Hout, M. (2020). The role of question characteristics in designing and evaluating survey questions. *Advances in Questionnaire Design, Development, Evaluation and Testing*, 117-152.

Fowler, F. J., & Cosenza, C. (2012). Writing effective questions. In *International handbook of survey methodology* (pp. 136-160). Routledge.

Giesen, D., Meertens, V., Vis-Visschers, R., & Beukenhorst, D. (2012). Questionnaire development. *Statistics Netherlands*, 12(3), 1-82.

Schaeffer, N. C., & Dykema, J. (2020). Advances in the science of asking questions. *Annual Review of Sociology*, 46(1), 37-60.

שאלון הסקר הוא כלי המדידה, אך הוא משמש גם לתקשורת בין עורך הסקר ובין המשיב, בעיקר כשאיסוף הנתונים נעשה ללא מראיין. לפיכך, על השאלון לכלול הסבר על מטרות הסקר ועל הנושאים שהוא בודק, על מבצע הסקר ובמידת האפשר גם על מזמין הסקר. כמו כן, עליו לכלול פרטים שקשורים לאתיקה, סודיות ופרטיות של הנדגם, למשל: האם השתתפות בסקר היא מבחירה או מכורח (כגון מכוח חוק או תקנה); אם נדרשת הסכמה מדעת; אם הנדגם הוא אנונימי ואם לא - כיצד יובטח חיסיון המידע.

רוב שאלות הסקר הן שאלות סגורות, אשר מאפשרות למשיב לבחור תשובה מבין מספר אפשרויות. שאלות סגורות יכולות להיות מסוגים שונים, למשל: שאלות דיכוטומיות (כגון "כן" / "לא", "נכון" / "לא נכון"), בחירת אפשרות אחת או יותר מתוך רשימה, בחירה של תדירות, עוצמה או מידה מתוך טווח תשובות נתון. על מנת שהשימוש בשאלות סגורות יהיה יעיל, עליהן להיות מנוסחות בקפידה ולעבור בדיקות איכות. לעתים נעשה שימוש בשאלות פתוחות, על מנת לאפשר למשיב לענות במילים שלו בהתאם לעולם הידע, העמדות והרגשות שלו. שימוש בשאלות פתוחות גוזל זמן, בהשבה ובקידוד, ועלול לגרום למידע חסר, אם למשיב אין מוטיבציה ורצון להשיב. עם זאת, שימוש בשאלות פתוחות עדיף כאשר נבדקות עמדות ותחושות שקשה לבטאן באמצעות סולם תשובות סגור וכאשר למחבר השאלון אין ודאות לגבי מיצוי אפשרויות התשובה (למשל: כיוון שלא ניתן לבצע בדיקות מקדימות).

- הערות כלליות להגברת איכות השאלון -

- יש להקפיד על הגדרה אופרציונלית מתאימה למושג שנמדד (למשל: ביקור רפואי עשוי להתייחס לביקור אצל מטפלים, רופאים או נותני שירות רפואי);
- בניסוח של שאלה, יש להתחשב בנושא השאלה (כגון: מצב בריאותי, עמדות פוליטיות ועוד), בסוג השאלה (למשל: אירועים והתנהגויות לעומת הערכות ושיפוטיות), בממד התשובה (למשל: התרחשות, תדירות, מֶשֶׁךְ, תזמון) ובמבנה השאלה (למשל: שימוש בשאלת סינון; צמידות שאלות בנושאים שונים אך עם ממד תשובה זהה).

- תכנון של סדר שאלות ומיקומן בשאלון -

- בהתאם לעניין, רצוי להשתמש בשאלות סינון, על מנת להימנע משאלות שאינן רלוונטיות לחלק מהמשיבים, דבר שעלול לגרום לתחושת נטל ולירידה באיכות ההשבה;
- רצוי לשמור על קשר לוגי בין שאלות סמוכות;
- רצוי לסדר את השאלות תוך התחשבות בהשפעה האפשרית של שאלות קודמות בשאלון על שאלות עוקבות. עם זאת, לעתים סדר השאלות נקבע בהתאם לשיקולים אחרים כגון: השוואה לסקרים אחרים בינלאומיים או משנים קודמות;
- השאלות בראשית השאלון צריכות להתייחס בצורה ברורה לנושא הסקר, לעורר עניין במשיב ולהיות קלות ונעימות להשבה;

- בדרך כלל, שאלות שעוסקות באותו נושא יקובצו יחד כאשר ראשית נשאלות שאלות כלליות ולאחריהן שאלות בנושאים ספציפיים;
- אם אפשר, יש למקם בסוף השאלון שאלות שעוסקות בנושאים רגישים ואשר עלולות לגרום למשיב לחוש חוסר נוחות.
- כתיבת הוראות והנחיות -
 - למראיינים צריך להיות ברור אילו חלקים מהסקר להקריא ואלו חלקים נועדו עבורם;
 - למראיינים צריך להיות קל להקריא את ההוראות והשאלות;
 - צריכה להיות אחידות בדרך הפנייה של המראיינים למשיבים, באופן הקראת השאלות ובמסירת הסברים ומידע מסייע.
- שימוש בשפה -
 - בניסוח של הוראות, שאלות ואפשרויות התשובה יש להשתמש במילים פשוטות ומוכרות ולהימנע, ככל הניתן, משימוש במונחים טכניים ובסלנג;
 - יש להשתמש בתחביר פשוט ובמשפטים קצרים, למשל: ללא ניסוח דקדוקי "סביל" ולהימנע מניסוח משפטים בשלילה ובשלילה כפולה;
 - יש להשתמש במונחים מוגדרים (ספציפיים) ומוחשיים (קונקרטיים) ולהימנע ממונחים כלליים ומופשטים;
 - יש להימנע מניסוח מעורפל, על מנת שהכתוב (הוראות, שאלות, אפשרויות תשובה) יפורש בצורה דומה על ידי כל המשיבים, בהתאם למשמעות הרצויה שכיוון אליה עורך הסקר.
- ניסוח שאלות -
 - יש לכלול בכל שאלה נושא אחד;
 - בשאלות על ידע יש להיערך למצב שלמשיב אין את הידע הנדרש לענות;
 - בשאלות על עמדה יש להיערך למצב שלמשיב אין עמדה בנושא שעליו הוא נשאל;
 - אם השאלה מתייחסת לתקופת זמן התקופה צריכה להיות מוגדרת היטב ורצוי שלא תהיה רחוקה מדי;
 - בשאלות שדורשות הערכה או חישוב יש להיערך למצב שלמשיב אין יכולת לבצע הערכה או חישוב;
 - יש להימנע משאלות שכוללות הנחות סמויות לגבי עמדות או התנהגות (למשל: על המשיב ו/או על מצבו; על יציבות תופעות או מצבים, שבפועל עשויים להשתנות);
 - יש להימנע מהצגת התנהגות או עמדה מקובלת, לפני שאלה על התנהגות או עמדה של המשיב, על מנת לצמצם רצייה חברתית.

- אפשרויות תשובה בשאלות סגורות -
 - יש להתאים את אפשרויות התשובה לסוג השאלה (למשל: שאלה על תדירות אינה יכולה להיות מלווה באפשרויות תשובה שמבטאות הסכמה / אי-הסכמה);
 - אפשרויות התשובה צריכות למצות את טווח המדידה הרלוונטי;
 - אפשרויות התשובה צריכות להיות ללא חפיפה ביניהן במשמעות, לפי סדר לוגי ותוך איזון בין קוטב חיובי לשלילי;
 - יש להימנע משימוש בפורמט תשובות "כן / לא", "מסכים / לא מסכים", "נכון / לא נכון", כיוון שסולמות אלה מזמנים הטיה לבחירת תשובות חיוביות (acquiescence);
 - עם זאת, במצב שניתן לבחור יותר מתשובה אחת מרשימה ארוכה, מומלץ לנסח כל אפשרות תשובה כשאלה נפרדת עם אפשרות השבה דיכוטומית (כדי להתגבר על הטיה לבחור אפשרויות תשובה ראשונות בהשבה עצמית, או אחרונות בריאיון);
 - יש להימנע מהבלטת אפשרות התשובה "לא יודע", כיוון שיש נטייה לבחור בה כשמשיב עלול שלא להשקיע מאמץ בהשבה.
- אפשרויות תשובה בסולמות דירוג (rating scales) לבדיקת תדירות ועוצמה -
 - רצוי להשתמש בחמש עד שבע אפשרויות תשובה (לא כולל תשובות מסוג: "לא יודע" או "לא רלוונטי");
 - ככל שהשאלה דורשת דיוק ו/או רגישות רבים יותר, מתאים להשתמש בטווח אפשרויות תשובה רחב יותר;
 - מומלץ להצמיד תווית מילולית לכל ספרה בסולם התשובות, על מנת ליצור אחידות בהבנה של כל אפשרות תשובה ולהגביר את מהימנות התשובות;
 - יש לבחור בסולם תשובות זוגי או אי-זוגי, בהתאם לצורך בקבלת עמדה ברורה מהמשיב או באפשרות להתלבט או להסס בתשובה.
- המלצות לאיסוף המידע -
 - יש לעצב את השאלון בהתאם לאופן איסוף הנתונים (למשל: מספר שאלות בכל דף, גודל גופנים, רווחים בין שאלות ועוד);
 - לרוב, מומלץ לשלב שיטות איסוף שונות במטרה להגביר שיעורי השבה;
 - בפנייה לנדגם להשתתפות בסקר יש לכלול הסבר על מטרות הסקר, על הנושאים שייבדקו, על מבצע הסקר ואם ניתן גם על מזמין הסקר, כדי להגביר אמון ושיתוף פעולה מצד הנדגם;
 - מומלץ לבחור בשיטת איסוף בדרך של פנים אל פנים כשחשוב להגיע לשיעור השבה גבוה וכשנדרש להגיע לקבוצות באוכלוסייה שנוטות שלא להשיב באופן אחר;
 - אם הסקר בודק נושאים רגישים אשר מעוררים רצייה חברתית, מומלץ לאפשר מענה עצמאי של הנדגם לסקר;

- אם נעשה שימוש במראיינים לצורך איסוף המידע עליהם להיות בעלי יכולת וכישורים שיתאימו לאוכלוסייה הנסקרת ולנושאים שנבדקים בסקר;
- אם נעשה שימוש בשאלון ממוחשב יש להתאים אותו למשתמשים (משיב ו/או מראיין), כך שהוא יהיה קל לניווט ולהתמצאות.

תכנון עיבוד וניתוח

הכנת הנתונים לעיבוד

במדגמים הסתברותיים, בשלב ראשון לאחר קבלת הנתונים שנאספו, יש להבחין בין מקרים השייכים (eligible) לאוכלוסיית הסקר, בין מקרים לא-שייכים (ineligible) (מכונים גם "מקרי אפס") ובין מקרים שלא ידוע אם הם שייכים או לא שייכים. מקרים לא-שייכים עשויים להיות עקב טעות בהכללת יחידות דגימה שאינן שייכות לאוכלוסייה בתוך מסגרת הדגימה, עקב אפיון שגוי של נדגמים במסגרת הדגימה (למשל: מקום מגורים לא מעודכן) ועוד. לאחר מכן, מתוך המקרים ששייכים יש לאפיין את מי שלא השתתפו בסקר לעומת אלו שהשתתפו.²¹

חוסר השתתפות בסקר

מבין הנדגמים שלא השתתפו בסקר יש לאפיין כל נדגם, בהתאם לסיבות אפשריות, להלן:

- כישלון ביצירת קשר (non-contact) – כולל מצבים שבהם עורכי הסקר לא הצליחו ליצור קשר (למשל: לא התקבל מענה לשיחות טלפון, לא התקבל מענה כתגובה לפנייה פנים או פנים של עורכי הסקר אל הנדגם, לא התקבלה עדות לכך שמכתב שנשלח הגיע, או לא הגיע, לכתובת וכדומה). במידת האפשר, יש לתעד בעת איסוף המידע אלו ניסיונות נעשו ליצירת קשר (למשל: האם הושארה הודעה קולית ו/או הודעת טקסט בטלפון, מתי הושארה הודעה ועוד);
- סרבנות (refusal) והפסקת מענה (break-off) – מבטאות עמדה אקטיבית של הנדגם, או של מרואיין חליפי, כלפי הסקר. המצב והברור ביותר הוא סירוב להיחקר (להתראיין, לענות על שאלון או להימדד בכל דרך אחרת) לאחר שעורכי הסקר יצרו קשר עם הנדגם או עם מרואיין חליפי. מצב דומה, אך פחות קיצוני הוא הפסקת מענה לאחר שהנדגם, או מרואיין חליפי, נתן את הסכמתו להשתתף והמענה לסקר התחיל. במידת האפשר, יש לתעד בעת איסוף המידע את הסיבות לסרבנות ולהפסקת ההשתתפות, על מנת לסייע באיתור ובהתמודדות עם הטיית אי-השבה;

²¹ להרחבה ראו –

American Association for Public Opinion Research (2016). Standard definitions: Final dispositions of case codes and outcome rates for surveys. *Encyclopedia of Survey Research Methods*.

- חוסר יכולת להשתתף מסיבות אחרות – משקף מקרים שבהם לא היה סירוב להשתתף בסקר אך בפועל הנדגם, או מרואיין חליפי, לא השתתף בסקר. למשל: מוות של נדגם (שהתרחש לאחר שהמדגם הוכן), לקות פיזית ו/או קוגניטיבית או אחרת שמונעת ממנו להשתתף בסקר (ולא ניתן לחקור מרואיין חליפי), בעיות שפה (הנדגם, או מרואיין חליפי, אינו מבין ו/או אינו דובר את שפת הסקר), בעיות תקשורת (למשל: בעיות קליטה בטלפון), חוסר נגישות של הנדגם לסקר (למשל: נמצא בנהיגה בעת פנייה אליו) ועוד.

אמות מידה להשתתפות בסקר

מבין כל מי שהשתתף בסקר חשוב להגדיר האם מענה לסקר (לשאלון, לריאיון) נעשה באופן מלא, באופן חלקי או שיש להגדירו כמענה חסר, שיש לסווגו כמענה שהופסק. ישנן סיבות שונות לכך שלא כל השאלות נענות, למשל: יש שאלות שאינן רלוונטיות לנדגם, יש שאלות שהנדגם, או מרואיין חליפי, אינו יכול או אינו רוצה לענות עליהן, יש שאלות שמסיבות אחרות לא נענו.

נוסף לאלה, יש שאלות שהן מרכזיות יותר עבור עורך הסקר לעומת שאלות אחרות שחשיבותן פחותה. לפיכך, חשוב להגדיר אלו שאלות רלוונטיות לכל נדגם ומה הן שאלת הליבה בסקר. ניתן להיעזר באמות המידה להלן לשם סיווג מידת המענה לשאלון הסקר (AAPOR, 2016²¹):

- מענה על אחוז מסוים משאלות הסקר שהן רלוונטיות לנדגם (applicable);
 - מענה על אחוז מסוים משאלות הליבה בסקר, אם הוגדרו. (למשל: שאלות שעוסקות במשתני התוצאה שהוגדרו במטרות הסקר ו/או במשתנים מסבירים.)
- להלן דוגמה להגדרה של מענה מלא, מענה חלקי והפסקת מענה (break-off) באמצעות שימוש באמות המידה לעיל:

- מענה על שאלות רלוונטיות –
מענה מלא = יותר מ-80%; מענה חלקי = 50% עד 80%; הפסקת מענה = פחות מ-50%;
או
מענה על שאלות ליבה –
מענה מלא = 100%; מענה חלקי = 50% עד 99%; הפסקת מענה = פחות מ-50%;
או
שילוב של הרשום לעיל, למשל -
מענה מלא = 100% מענה על שאלות הליבה ו-80% מענה על שאר השאלות.

יש להדגיש כי אלו דוגמאות בלבד וכי על מזמין הסקר ו/או מבצע הסקר להגדיר באופן ברור את סוגי ההשתתפות בסקר, את אמות המידה ואת המשקלות שיקבעו.

הטיית אי-השבה מתרחשת כשיעור ההשבה נמוך, כשקיים הבדל משמעותי בין מאפייני המשיבים לעומת הלא-משיבים וכאשר שקיים קשר בין הנטייה להשיב ובין התשובות למשתנה הנחקר. הטיית אי-השבה אינה מתקיימת כשהגורמים להסתברות לאי-השבה אינם תלויים בתגובה למשתנה הנחקר (מצב של תצפיות חסרות באקראי באופן מלא - missing completely at random). במצב זה, ובהנחה שאין הטיית אחרות, כל ניסיון להגדיל את שיעור ההשבה עלול להעלות, ללא צורך, את עלות הסקר ולהוריד את איכות המידע (למשל, על ידי גיוס של יותר משיבים שיענו רק על חלק מהשאלון, כלומר, הגברת מקרים של אי-השבה לפריט).

לרוב, קיים קשר בין ההסתברות להשיב ובין סוג התגובה למשתנה הנחקר, כך שקיים גורם משותף (common cause) אחד, או מספר גורמים, הן להסתברות להשיב והן לתגובה למשתנה הנחקר. דוגמה לכך היא עניין מיוחד שיש לחלק מהמשיבים כלפי נושא שנמדד בסקר, למשל, בנייה של מרכז קניות ליד ישוב כפרי. במקרה כזה, סביר שהנושא משמעותי מאוד עבור תושבים שגרים באותו ישוב כפרי ופחות משמעותי עבור תושבים באזורים אחרים. לפיכך, סביר ששיעור ההשבה של תושבי היישוב יהיה גבוה יותר, לעומת שיעור השבה באזורים אחרים. מצב זה עשוי להתבטא למשל בשיעורים שונים של סרבנות להשתתף בסקר (refusal) ואולי אף בשיעורים שונים של נשירה (break-off), כלומר, הפסקת ההשבה במהלך מילוי השאלון. בנוסף, סביר שהנושא ישפיע על עמדותיהם (באופן חיובי או שלילי) בצורה שונה לעומת המצב אל תושבי אזורים אחרים. אם גורמים משותפים אלה נמדדים במחקר (כמו למשל, הישוב שבו המשיב מתגורר), ניתן, בעזרת התאמות סטטיסטיות, להתגבר על השפעתם ועל הטיית אי-השבה. כך, בתוך קבוצות הומוגניות (של בעלי עניין וחוסר עניין כלפי הנושא הנחקר), לא תהיה הטיית אי-השבה. כיוון שקיימת הטיית שיטתית בין הקבוצות, בתוך כל קבוצה התצפיות הקיימות יתפלגו בצורה דומה לתצפיות החסרות; זהו מצב של תצפיות חסרות באקראי (missing at random).

²² להרחבה ראו:

Čehovin, G., Bosnjak, M., & Lozar Manfreda, K. (2022). Item Nonresponse in Web Versus Other Survey Modes: A Systematic Review and Meta-Analysis. *Social Science Computer Review*, 08944393211056229.

Felderer, B., Kirchner, A., & Kreuter, F. (2019). The effect of survey mode on data quality: Disentangling nonresponse and measurement error bias. *Journal of Official Statistics*, 35(1), 93-115.

Groves, R. M. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly*, 70, 646–75.

Tourangeau, R. (2020). How errors cumulate: Two examples. *Journal of Survey Statistics and Methodology*, 8(3), 413-432.

מצב נוסף של הטיית אי-השבה הוא כאשר הגורם להסתברות להשבה נובע ממשנתה החקירה עצמו (survey variable cause). לדוגמה, אם סקר מבקש לבדוק שליטה בשפה העברית וחלק מהמשיבים אינם דוברי עברית, הרי שמידת השליטה בשפה תקבע את ההסתברות להשבה. זהו מצב של תצפיות חסרות שלא באקראי (not missing at random). דוגמה נוספת היא מצב שבו שאלון עוסק בנושאים רגישים וחלק מהמשיבים מסרבים להשתתף בסקר כיוון שאינם מעוניינים לדווח על התנהגויות לא נורמטיביות או שאינן רצויות מבחינה חברתית. המחקר מראה כי אם הנדגמים בכל זאת משתתפים בסקר, הם ייטו לתת דיווח שקרי. כמו כן, שילוב של הטיית ייצוג והשבה ושל הטיית מדידה גורם להטיית אומדני הסקר, בעיקר כשהם מופקים ממדגמים אינטרנטיים לא-הסתברותיים.

מדדי תפוקה (outcome rates) למדגמים הסתברותיים

לאחר שיוגדרו מצבים של אי-השתתפות ושל השתתפות בסקר ולאחר שתיבדק איכות ההשבה, ניתן לחשב מספר מדדי תפוקה: שיעור השבה (response rate); שיעור שיתוף פעולה (cooperation rate); שיעור סרבנות (refusal rate); שיעור יצירת קשר (contact rate). אלו יתוארו להלן. מקובל לחשב את מדדי התפוקה באמצעות רכיבים אלה (AAPOR, 2016²¹):

- מספר שאלוני סקר עם מענה מלא (I – complete interview);
- מספר שאלוני סקר עם מענה חלקי (P – partial interview);
- מספר מקרי הסרבנות ו/או עם מענה שהופסק (R – refusal and break-off);
- מספר מקרי חוסר קשר (NC – non-contact);
- מספר מקרי אי-השתתפות אחרים (שאינם סרבנות או מענה שהופסק) (O – other);
- מספר לא ידועים אחרים (UH – unknown if household / occupied);
- (UO – unknown, other);
- הערכת יחס מקרים שייכים מתוך מקרים שלא ידוע אם הם שייכים (e).

יש להצדיק את בחירת הרכיבים בכל אחד מהמדדים שיתוארו בהמשך (למשל: מספר שאלוני סקר עם מענה מלא או חלקי).

שיעור השבה (response rate)

שיעור השבה מבטא את היחס בין מספר המקרים שבהם התקבל מענה לשאלונים ובין מספר המקרים הכללי במדגם. ניתן לחשב שישה סוגים של שיעורי השבה, הראשון שבהם (RR1) מייצג את שיעור ההשבה המזערי והאחרון (RR6) את שיעור ההשבה המרבי – RR1 – יחס בין מספר שאלונים עם מענה מלא ובין כלל המקרים במדגם;

$$RR1 = \frac{I}{(I + P) + (R + NC + O) + (UH + UO)}$$

RR2 – יחס בין מספר שאלונים עם מענה מלא וחלקי ובין כלל המקרים במדגם;

$$RR2 = \frac{(I + P)}{(I + P) + (R + NC + O) + (UH + UO)}$$

RR3 – יחס בין מספר שאלונים עם מענה מלא ובין כלל המקרים במדגם, כולל הערכת יחס מקרים שייכים מתוך מקרים שלא ידוע אם הם שייכים;

$$RR3 = \frac{I}{(I + P) + (R + NC + O) + e(UH + UO)}$$

RR4 – יחס בין מספר שאלונים עם מענה מלא וחלקי ובין כלל המקרים במדגם, כולל הערכת יחס מקרים שייכים מתוך מקרים שלא ידוע אם הם שייכים;

$$RR4 = \frac{(I + P)}{(I + P) + (R + NC + O) + e(UH + UO)}$$

RR5 – יחס בין מספר שאלונים עם מענה מלא ובין כלל המקרים במדגם, בהנחה שאין מקרים שייכים בתוך הלא ידועים, או שאין מקרים לא ידועים;

$$RR5 = \frac{I}{(I + P) + (R + NC + O)}$$

RR6 – יחס בין מספר שאלונים עם מענה מלא וחלקי ובין כלל המקרים במדגם, עם אותה הנחה לעיל לגבי מקרים שייכים בתוך מקרים לא ידועים.

$$RR6 = \frac{(I + P)}{(I + P) + (R + NC + O)}$$

שיעור שיתוף פעולה (cooperation rate)

שיעור שיתוף פעולה מבטא את היחס בין כל המקרים שנחקרו (ענו על שאלון, תושאלו על ידי סוקר וכו') ובין כלל המקרים השייכים, שאי פעם נוצר איתם קשר. ניתן לחשב ארבעה סוגים של שיעורי שיתוף פעולה –

COOP1 – יחס בין מספר שאלונים עם מענה מלא ובין כלל המקרים השייכים שנוצר עימם קשר;

$$COOP1 = \frac{I}{(I + P) + R + O}$$

COOP2 – יחס בין מספר שאלונים עם מענה מלא וחלקי ובין כלל המקרים השייכים שנוצר עימם קשר;

$$COOP2 = \frac{(I + P)}{(I + P) + R + O}$$

COOP3 – יחס בין מספר שאלונים עם מענה מלא ובין כלל המקרים השייכים שנוצר עימם קשר, להוציא מקרים שאין ביכולתם להשתתף;

$$COOP3 = \frac{I}{(I + P) + R}$$

COOP4 – יחס בין מספר שאלונים עם מענה מלא וחלקי ובין כלל המקרים השייכים שנוצר עימם קשר, להוציא מקרים שאין ביכולתם להשתתף.

$$COOP4 = \frac{(I + P)}{(I + P) + R}$$

שיעור סרבנות (refusal rate)

שיעור סרבנות מבטא את היחס בין כל המקרים שבהם היה סירוב לענות לסקר או שהייתה הפסקת מענה לסקר (break-off) ובין כלל המקרים במדגם. ניתן לחשב שלושה סוגים של שיעורי סרבנות – REF1 – יחס בין מספר המקרים שבהם היה סירוב או הפסקת מענה ובין כלל המקרים במדגם;

$$REF1 = \frac{R}{(I + P) + (R + NC + O) + (UH + UO)}$$

REF2 – יחס בין מספר המקרים שבהם היה סירוב או הפסקת מענה ובין כלל המקרים במדגם, כולל הערכת יחס מקרים שייכים מתוך מקרים שלא ידוע אם הם שייכים;

$$REF2 = \frac{R}{(I + P) + (R + NC + O) + e(UH + UO)}$$

REF3 – יחס בין מספר המקרים שבהם היה סירוב או הפסקת מענה ובין כלל המקרים במדגם, בהנחה שאין מקרים שייכים בתוך הלא ידועים, או שאין מקרים לא ידועים.

$$REF3 = \frac{R}{(I + P) + (R + NC + O)}$$

שיעור יצירת קשר (contact rate)

שיעור יצירת קשר מבטא את היחס בין כל המקרים שבהם נוצר קשר עם בן משק בית אחראי ובין כלל המקרים במדגם. ניתן לחשב שלושה סוגים של שיעורי קשר – CON1 – יחס בין מספר שאלונים עם מענה מלא וחלקי, מקרים שבהם היה סירוב או הפסקת מענה ומקרים אחרים של אי-השתתפות ובין כלל המקרים במדגם;

$$CON1 = \frac{(I + P) + R + O}{(I + P) + R + O + NC + (UH + UO)}$$

CON2 – יחס בין מספר שאלונים עם מענה מלא וחלקי, מקרים שבהם היה סירוב או הפסקת מענה ומקרים אחרים של אי-השתתפות ובין כלל המקרים במדגם, כולל הערכת יחס מקרים שייכים מתוך מקרים שלא ידוע אם הם שייכים;

$$CON2 = \frac{(I + P) + R + O}{(I + P) + R + O + NC + e(UH + UO)}$$

CON3 – יחס בין מספר שאלונים עם מענה מלא וחלקי, מקרים שבהם היה סירוב או הפסקת מענה ומקרים אחרים של אי-השתתפות ובין מקרים שייכים במדגם.

$$CON3 = \frac{(I + P) + R + O}{(I + P) + R + O + NC}$$

מדדי תפוקה בפנלים הסתברותיים ולא-הסתברותיים²³

בהמשך לאמור בנספח זה בחלק שעוסק ב"דגימה מתוך פנלים", ניתן לחשב מדדי תפוקה נוספים שמתאימים להשתתפות בפנלים (אינטרנטיים ולא-אינטרנטיים):

- שיעור גיוס (recruitment rate), מתאימם לפנל הסתברותי בלבד;
- שיעור השלמת פרופיל (profile rate);
- שיעור השלמת שאלון (completion rate);
- שיעור השבה כולל (cumulative response rate).

שיעור גיוס (recruitment rate)

שיעור גיוס (RECR) מבטא את היחס בין מספר המקרים שבהם התקבלה הסכמה ראשונית (initial consent – IC) להשתתף בפנל ובין מספר המקרים הכללי במדגם. לתשומת לב, שיעור גיוס ניתן לחשב רק עבור פנלים הסתברותיים.

$$RECR = \frac{IC}{IC + (R + NC + O) + e(UH + UO)}$$

²³ Callegaro, M., & DiSogra, C. (2008). Computing response metrics for online panels. *Public opinion quarterly*, 72(5), 1008-1032.

שיעור השלמת פרופיל (profile rate)

במקרים רבים חברות בפנל נחשבת רק לאחר רישום פרטי הפרופיל. שיעור השלמת פרופיל (PROR) מבטא את היחס בין מספר המקרים שמלאו פרטים באופן מלא (I) או חלקי (P) ובין כלל המקרים שנתנו את הסכמתם להשתתף בפנל. שיעור השלמת פרופיל מתאים לפנלים הסתברותיים ולפנלים לא-הסתברותיים. מכיוון שהמענה לשאלון פרטי הפרופיל עלול להיות חלקי (לכלול רק חלק מהפרטים), ניתן לחשב שני שיעורי השלמת פרופיל –

$$\text{PROR1} = \frac{I}{(I + P) + (R + NC + O)}$$

$$\text{PROR2} = \frac{(I + P)}{(I + P) + (R + NC + O)}$$

שיעור השלמת שאלון (completion rate)

שיעור השלמת שאלון (COMR) מבטא את היחס בין מספר המקרים שבהם התקבל מענה לשאלונים של סקר מסוים ובין מספר חברי הפנל שמלאו שאלון פרופיל. ניתן לחשב שני סוגים של שיעורי השלמת שאלון –

COMR1 – יחס בין מספר שאלונים עם מענה מלא ובין כלל המקרים בפנל שמלאו פרטי פרופיל;

$$\text{COMR1} = \frac{I}{(I + P) + (R + NC + O)}$$

COMR2 – יחס בין מספר שאלונים עם מענה חלקי ובין כלל המקרים בפנל שמלאו פרטי פרופיל;

$$\text{COMR2} = \frac{(I + P)}{(I + P) + (R + NC + O)}$$

שיעור השבה כולל (cumulative response rate)

שיעור השבה כולל (CUMRR) הוא מכפלה של שלושת המדדים שתוארו לעיל:

$$\text{CUMRR} = \text{RECR} \times \text{PROR} \times \text{COMR}$$

יש לציין כי בפועל, המדד עשוי לכלול רכיבים נוספים. ראשית, הגיוס לפנל עשוי להיות על בסיס קבוע ורציף, כך שההרכב שלו משתנה עם חלוף הזמן. לכן, שיעור הגיוס עשוי להיות מדד המורכב מכמה תת-מדדים. בנוסף, לעיתים נעשה שימוש בשאלוני סינון נוסף לשאלון הפרופיל, על מנת לקבוע את התאמתו של חבר בפנל לסקר מסוים, מרכיב שיש להוסיף לחישוב שיעור ההשבה. לבסוף, אם הסקר שנערך הוא סקר אורך (longitudinal) או שהוא נערך במספר גלים, עלולה להיות נשירה מהפנל, דבר שמוסיף רכיב נוסף לחישוב (להרחבה ראו, Callegaro & DiSogra, 2008²³).

שיטות להתאמה ותיקון במדגמים לא-הסתברותיים²⁴

לרוב, חוקרים מתקפים את ממצאי הסקר הלא-הסתברותי, באמצעות השוואה לממצאים של סקר הסתברותי, למשל, באמצעות השוואת ממוצעים בין שני המדגמים (כאשר מדובר בסולמות מדרג), בהתפלגויות אחוזים, ואף בדפוסים של קשרים בין משתנים ושל הבדלים בין קבוצות. מומלץ להיזהר במינוח, אם מדגם לא-הסתברותי משמש להסקה על האוכלוסייה. שימוש במונחים הסתברותיים "סטנדרטיים" (כגון "טווח ביטחון"), עלולים להטעות ולהטות את הקורא ולגרום לו

²⁴ להרחבה ראו:

Cornesse et al. (2020). A Review of Conceptual Approaches and Empirical Evidence on Probability and Nonprobability Sample Survey Research. *Journal of Survey Statistics and Methodology*, 8, 4–36.

Elliott, M. R., & Valliant, R. (2017). Inference for nonprobability samples. *Statistical Science*, 32(2), 249-264.

Mercer, A. W., Kreuter, F., Keeter, S., & Stuart, E. A. (2017). Theory and practice in nonprobability surveys: parallels between causal inference and survey inference. *Public Opinion Quarterly*, 81(S1), 250-271.

Pasek, J., & Krosnick, J. A. (2020). Relations Between Variables and Trends Over Time in RDD Telephone and Nonprobability Sample Internet Surveys. *Journal of Survey Statistics and Methodology*.

Scherpenzeel, A. C., and J. G. Bethlehem (2011). How representative are online panels? Problems of coverage and selection and possible solutions. In M. Das, P. Ester, & L. Kaczmarek (Eds.), *Social and behavioral research and the Internet: Advances in applied methods and research strategies* (pp. 105–132). New York: Routledge.

Vehovar, V., Toepoel, V., & Steinmetz, S. (2016). Non-probability sampling. *The Sage handbook of survey methods*, 329-345.

להאמין שמדובר בערכים אמיתיים באוכלוסייה, למרות שבפועל ערכים אלה עלולים להיות שונים והסטייה מהם עלולה להיות גדולה יותר מזו שמדווח עליה.

להלן, יתוארו גישות לתיקון ולהתאמות שניתן לנקוט בהן בניתוח. גישות אלה מבוססות על שיטות לכיול גלובלי (global adjustment) ועל שיטות לכיול מותאם-תוצאה (outcome-specific adjustment).

שיטות לכיול גלובלי (global adjustment)

שיטות לכיול גלובלי הן שיטות כיול ששאלות מהדגימה ההסתברותית. שיטות אלה משתמשות במודל כדי ליצור התאמה אחת שניתן ליישם אותה בכל ניתוח, ללא קשר לתוצאה המבוקשת. ההנחה היא שברגע שמתקנים את המקורות הידועים להטיה, כל שאר ההטיות יהיו בלתי תלויות בסיכוי של יחידה מסוימת באוכלוסייה להשתתף בסקר ולכן כל יחידה שנדגמה בתת-קבוצת מסוימת באוכלוסייה תהיה דומה ליחידות שלא נדגמו באותה תת-קבוצה, בכל הנוגע לכל היסק שייעשה מהנתונים. ההטיה תהיה קטנה יותר, בהינתן התנאים האלה: (1) הנטייה להשתתף בסקר דומה בתוך בכל תת-קבוצה מוגדרת, (2) ערכי משתנה התוצאה (החקירה) דומים בתוך כל תת-קבוצה, (3) אין קשר בין הנטייה להשיב ובין משתנה התוצאה בכל תת-קבוצה. ניתן לעמוד בתנאים אלה אם קיים קשר חזק בין משתנה התוצאה למשתני הריבוד, שיוצרים את תת-הקבוצות.

אחת הגישות הנפוצות, מבין שיטות הכיול הגלובלי, היא שקלול מתוקן (calibration weighting) שבו ערכי סך הכול של המדגם, במאפיינים מסוימים, מתוקנים בהתאם לערכים באוכלוסייה. הערכים על האוכלוסייה אמורים להגיע ממקור מידע איכותי כגון מידע מפקדי או סטטיסטיקה רשמית. ההנחה שקיימת בשימוש במשקלות גלובליים היא, שברגע שמתקנים את המקור להטיה (כלומר, אותם משתנים שזוהו ככלל המשתנים המתערבים שגורמים להטיה), כל הטיות הבחירה ינוטרלו. לפיכך, אסטרטגיה זו מניחה שכל היסק שייעשה לגבי יחידות שנדגמו בתת-קבוצות מסוימות באוכלוסייה יהיה תקף גם ליחידות שלא נדגמו בתת-קבוצות אלה.

גישה נוספת היא שקלול באמצעות ציוני נטייה (propensity scores) בשלב הניתוח. הרעיון הבסיסי הוא לחשב את ההסתברות להשיב (או לא להשיב) לסקר הלא-הסתברותי, באמצעות שימוש במקור מידע השוואתי, שהינו לרוב מדגם הסתברותי גדול דיו. לרוב, ציוני הנטייה מופקים באמצעות רגרסיה לוגיסטית, תוך שימוש במשתנים מסבירים שונים (רקע, עמדה וכדומה), שנאספים הן במדגם ההסתברותי והן במדגם הלא-הסתברותי. על מנת להשיג חוסר הטיה, יש להשתמש במשתני בקרה מתאימים. בנוסף, נדרש לדעת מהי ההתפלגות של המשתנים הללו באוכלוסייה (מידע שלרוב קשה להשגה). לאחר איסוף המידע באמצעות מדגם לא-הסתברותי, משתמשים ציוני הנטייה (המשקלות) לכל הניתוחים שייעשו על המידע.

שיטה שלישית, אמידה דמוית מערך הסתברותי (pseudo design-based estimation) משלבת בעת הניתוח מידע מהמדגם הלא-הסתברותי עם מידע מהמדגם ההסתברותי. ראשית, נאמדות

הסתברויות להכללת היחידות במדגם הלא-הסתברותי (pseudo inclusion probabilities) על בסיס מערך משתנים שמשותפים לשני המדגמים. שנית, לאחר הצמדת ההסתברויות ליחידות אלה, ניתן להתייחס אליהן כאילו הן נבחרו באמצעות אותו המכניזם שהנחה את בחירת היחידות במדגם ההסתברותי. בשלב זה ניתן לנתח את שני המדגמים יחדיו.

כל שיטות הכיול הגלובלי מניחות שהמשתנים שמשמשים להתאמה, שמקורם במפקד או במדגם הסתברותי, הם ללא הטיה ונמצאים במתאם גבוה עם משתני המטרה ועם ההסתברות להיכלל במדגם הלא-הסתברותי. על מנת להגביר את התוקף של הנחות אלה יש צורך במקור נתונים מגוון שכולל מערך משתנים נרחב. עם זאת, יש לזכור כי שיטות גלובליות לשקלול עשויות לשפר אומדנים בחלק מהמקרים אך לא באחרים, וכי אין ביטחון שההטיות במדגם הלא-הסתברותי יבוטלו לחלוטין.

שיטות לכיול מותאם-תוצאה (outcome-specific adjustment)

שיטות לכיול מותאם-תוצאה מנסות להתאים מודל ספציפי למשתנה תוצאה (חקירה) מסוים שנבדק. בשיטות אלה נעשה ניסיון לשלוט במשתנים שקשורים לתהליך הבחירה של המשיבים, ואשר נמצאים במתאם עם משתנה התוצאה.

על פי גישה נוספת, תיקון נתמך מודל (model-assisted calibration), יצירת המשקלות המתוקנים נעשית באמצעות שימוש במודל שמנבא את ערכי המשתנים שבניתוח. הנחה מרכזית בגישה היא שהמודל מסוגל לנבא באופן מהימן ערכים במדגמים אחרים שלקוחים מאותה אוכלוסייה. הנחה נוספת היא שכל הפרמטרים הרלוונטיים באוכלוסיית-העל נכללים במודל. אחד המודלים שנעשה בהם שימוש הוא LASSO (Least Angle Shrinkage and Selection Operator), המאפשר הן בחירת משתנים למודל והן אמידה של הפרמטר; הוא גם מפיק תוצאות טובות יותר, לעומת מודלים אחרים, כשקיים מתאם גבוה (multicollinearity) בין משתני הבקרה.

גישה אחרת לכיול מותאם-תוצאה היא באמצעות רגרסיה היררכית וריבוד בדיעבד (post-stratification), שבה הרמות השונות ברגרסיה מאפשרות לשלב מספר רב של משתני בקרה ו/או אינטראקציות מסדר גבוה במודל הניבוי. המודל משמש כדי לאמוד את הערך הממוצע במספר רב של תאים מרובדים. בשונה משיטות פשוטות יותר (כגון raking), בהן די לדעת מהן ההתפלגויות השוליות, בשיטה זו יש צורך לדעת מהו הגודל היחסי של כל תא באוכלוסייה, או להעריך אותו באופן מהימן ממקור מידע חיצוני, כגון מידע מפקדי או מרשם אוכלוסין.

ככלל, הגישות שמשתמשות במדגמים לא-הסתברותיים כדי להסיק על אוכלוסייה, תלויות בשילוב של הנחות לגבי זניחות ההטיות ונגישות של מידע על הגורמים לבחירת המשיבים. בהנחות אלה ניתן להשתמש כדי לערוך התאמות לנתונים. למרות שהנחות אלה אינן שונות במהותן מההנחות שנמצאות בבסיס הדגימה ההסתברותית, קשה יותר להסתמך עליהן כיוון שיש לנו ידע מועט לגבי הגורמים שמובילים אנשים להשתתף במדגמים לא-הסתברותיים.

כפי שתואר לעיל, אחד מהמדדים לבדיקת התוקף של כלי המדידה היה תוקף תלוי קריטריון (criterion related validity), אשר מבטא את הקשר בין מידע שהופק ממדד בשאלון ובין נתונים עובדתיים שניתן לסמוך שהם מהימנים ותקפים. למעשה, ניתן לבדוק האם הממצאים משקפים את התופעה שהם אמורים לבדוק באמצעות בדיקת ההלימה לנתוני קריטריון (ראו לעיל, "תוקף תלוי קריטריון"). מקורות אפשריים עשויים להיות, למשל: נתוני סקר קודם, נתונים מנהליים וכיו"ב. לרוב, לא קיימים נתונים שמאפשרים להשוות אליהם את ממצאי הסקר ולכן יש לערוך בדיקות שישקפו את תוקף הממצאים, כלומר, ממצאים שנותנים ביטוי אמיתי וללא הטיה של התופעה שנבדקת. שני מדדים מרכזיים לבדיקת תוקף הממצאים הם תוקף חיצוני (external validity) ותוקף פנימי (internal validity).

תוקף חיצוני (external validity)

תוקף חיצוני מבטא את המידה שבה ממצאי הסקר ניתנים להכללה (generalizability) מעבר לכלל אוכלוסיית הסקר. לעתים מעוניינים להכליל את ממצאי הסקר לאוכלוסייה אחרת (transportability), דבר שדורש מאמץ אמפירי נוסף, כיוון שהמדגם אינו מהווה חלק מאוכלוסיית היעד. נהוג לבחון את הכללת ממצאי הסקר ביחס ליחידות, לסביבה, לזמן, ולאופן מדידת המשתנים:

- יחידות – המידה שניתן להכליל את ממצאי הסקר לכלל האוכלוסייה (למשל: דמיון בין מאפייני המשתתפים בסקר ובין מאפייני האוכלוסייה, דמיון בין מאפייני המשיבים והלא-משיבים);
- סביבה – המידה שניתן להכליל את ממצאי הסקר לסביבות רלוונטיות אפשריות אחרות, שונות מהסביבה שבה נאספו הממצאים (למשל: בסקר בריאות סביבות שונות עשויות לכלול מעבדה, בית-חולים, קופת חולים, אזורים במדינה ועוד);
- זמן – המידה שניתן להכליל את ממצאי הסקר לזמנים אחרים, שונים מהזמן שבו נאספו הממצאים (למשל: שעם חלוף הזמן לא יחולו שינויים באוכלוסייה);

²⁵ להרחבה ראו:

Findley, M. G., Kikuta, K., & Denly, M. (2021). External validity. *Annual Review of Political Science*, 24, 365-393.

Lavrakas, P. J., Kennedy, C., de Leeuw, E. D., West, B. T., Holbrook, A. L., & Traugott, M. W. (2019). Probability Survey-Based Experimentation and the Balancing of Internal and External Validity Concerns. In P.Lavrakas, M.Traugott, C.Kennedy, A.Holbrook, E.de Leeuw, & B.West (Eds.), *Experimental Methods in Survey Research* (pp. 1–18). John Wiley & Sons, Ltd.

Russell, M. (2022). Clarifying the Terminology of Validity and the Investigative Stages of Validation. *Educational Measurement: Issues and Practice*, 41(2), 25-35.

- הגדרת משתנים – המידה שבה ניתן להכליל את ממצאי הסקר שהושגו במדידה עם הגדרות אופרציונליות ("תרגום" של המושגים שנבדקים בסקר באמצעות מדידה תצפיתית) נתונות למדידה עם הגדרות אופרציונליות אחרות (למשל: טיפול בחרדה על ידי טיפול פסיכותרפי לעומת טיפול תרופתי; מדידת חרדה באמצעות דופק והזעה לעומת דיווח בשאלון; מדידת משתנים רלוונטיים אחרים שעשויים לתווך את השפעת הטיפול על התוצאות).
- תוקף חיצוני לקוי קשור להטיות ייצוג (כיסוי, דגימה ואי-השבה) ולהטיות בעיבוד הנתונים ובפעולות ההתאמה שנעשות לנתונים. ככל שמדגם מייצג את האוכלוסייה טוב יותר ויש בו פחות הטיות אי-השבה, כך יגבר התוקף החיצוני שלו.
- הוכחת תוקף חיצוני קשה יותר בסקר לא-הסתברותי ומומלץ לשים לב להיבטים האלה:
- אפשרות לקשר בין מאפייני הנדגמים ומשתנה החקירה – למשל: אם ניסוי בודק התנהגות של מבוגרים, ייתכן שמי שישתתף בסקר אינטרנטי לא ייצג היטב את כלל המבוגרים, עקב נגישות שונה לאינטרנט, הבדלים במיומנויות בשימוש במחשב ומאפיינים כלכליים-חברתיים;
- שימוש במשתני רקע עיקריים בלבד (מגדר, גיל, דת) אינו מספיק – יש לשקלל גם בהתאם למשתנים שקשורים לנטייה להשתתף בסקר ולמשתני החקירה (ראו עוד בנספח: "שיטות להתאמה ותיקון בשלב הדגימה", "שיטות להתאמה ותיקון במדגמים לא-הסתברותיים");
- שימוש בניתוח מבוסס מודל בדיווח על אומדנים ובדיקת השערות²⁶ – שימוש במודל סטטיסטי שיבטיח שניתן להתייחס למדגם הלא הסתברותי כאל מדגם מקרי פשוט;
- תשומת לב לקבוצות קטנות שנדרש לייצג אותן ולדווח עליהן בממצאים, מבחינת הטיות ייצוג ואפשרות להתמודדות איתן בעיבוד הנתונים;
- דוגמאות לשחזור של ממצאי הסקר במדגמים אחרים, רצוי הסתברותיים.

תוקף פנימי (internal validity)

תוקף פנימי מבטא את המידה שבה קיים קשר סיבתי בין המשתנים הבלתי תלויים בסקר ובין המשתנה התלוי.

על מנת להוכיח תוקף פנימי נדרשים שלושה תנאים:

- מתאם משמעותי בין המשתנה הבלתי תלוי (הגורם, המסביר, המנבא) למשתנה התלוי (התוצאה, המוסבר, המנובא);
- סדר זמנים שבו המשתנה הבלתי תלוי מתרחש לפני המשתנה התלוי;

²⁶ להרחבה ראה:

Sterba, S. K. (2009). Alternative model-based and design-based frameworks for inference from samples to populations: From polarization to integration. *Multivariate behavioral research*, 44(6), 711-740.

- יכולת להפריך הסברים חלופיים לכך ששינוי במשתנה הבלתי תלוי הוא זה שגורם לשינוי במשתנה התלוי. תנאי זה מושג על ידי תפעול של מערך ניסויי מבוקר, שכולל מערך בדיקות "לפני-אחרי", עם הקצאה מקרית של נדגמים לקבוצות ניסוי וביקורת.

בהוכחת תוקף פנימי החוקר אמור להתמודד עם איומים שונים, כשהמרכזיים ביניהם הם אלה:

- היסטוריה (history) – אירועים חיצוניים שמתרחשים בזמן המחקר, אשר עלולים להשפיע על תנאי הטיפול או על התופעה (התוצאה) שנבדקת;
- בשילה (maturation) – שינויים פנימיים שמתרחשים במשתתפים במהלך המחקר (למשל: עייפות, התבגרות, רכישת ניסיון) ועלולים להשפיע על התוצאה שנבדקת;
- היבחנות (testing) – כל מדידה שנעשית על משתתפי המחקר עשויה להשפיע על תגובתם למדידה חוזרת (למשל: עקב רכישת ניסיון, זכירת שאלות ותשובות ועוד);
- מכשור (instrumentation) - בין המדידות עלולים להתרחש שינויים במכשיר המדידה או בתפעול שלו (למשל: שינוי בנוסח של שאלה ו/או בדרך הצגתה על ידי מראיין), דבר שעלול להשפיע על תגובה למשתנה הבלתי תלוי ו/או על מדידת המשתנה התלוי;
- ברירה או הצבה (selection) – שיוך נבדקים לקבוצות ניסוי וביקורת ללא הקצאה מקרית עלולה לגרום להבדלים בין הקבוצות באופן התגובה לטיפול ו/או בעוצמת ההשפעה של הטיפול על התוצאה;
- נשירה (mortality / attrition) – הפסקת השתתפות במחקר של משתתפים בקבוצות הניסוי והביקורת ללא שליטה או בקרה על שיעורי הנושרים מכל קבוצה ועל מאפייניהם.

אתיקה

כל פעולה סטטיסטית (כגון איסוף, עיבוד, ניתוח של מידע) אודות אנשים או גופים, צריך להתנהל בהתאם לעקרונות אתיים מקובלים. שמירה על אתיקה חשובה מסיבות רבות, למשל: להגן על זכויותיהם וכבודם של המשתתפים בסקר ושל קבוצות האוכלוסייה שמוצגות בו; לכבד את זכויותיהם של בעלי עניין המעורבים, לרבות מזמיני הפעולה הסטטיסטית ואלה שמיועדים להשתמש בנתונים ובממצאים; להגביר את האמון בגופים ממשלתיים, ציבוריים, אקדמיים, מסחריים ואחרים שעורכים מחקרים וסקרים וכן להגביר את הפיקוח ואת הבקרה על גופים אלה.

ההנחיות בנושא אתיקה מיועדות לכל מי שמתכנן סקרים או עורך אותם, לכל מי שקשור לטיפול במידע ולשמירתו וכן לגופים מומחים שקשורים לכך. האחריות ליישום של הנחיות האתיקה היא בעיקר של מזמין הפעולה הסטטיסטית, של המבצע אותה ושל מי ששומר על הנתונים.

הצורך לשמור על אתיקה בביצוע סקרים ומחקרים אינו חדש, אך הוא מקבל דגש לאור השינויים שמתרחשים בפרקטיקה של ביצוע הסקרים, לרבות שימוש במקורות מידע חדשים ובטכנולוגיות חדשות לייצור מידע, כמו גם הצורך להגן טוב יותר על פרטיותם של הנחקרים ועל סודיות הנתונים (United Nations, Economic Commission for Europe Conference of European Statisticians, 2024)²⁷. לפיכך, שמירה על אתיקה בסקרים, או בכל פעולה סטטיסטית שהיא, היא משימה נדרשת וחשובה.

פרטיות וסודיות

סיכון לפגיעה בפרטיות

ישנם מקרים שהסקר, או פעולה סטטיסטית אחרת (למשל: ניתוח של נתונים מנהליים), כוללים טיפול במידע מזהה, או כזה שמאפשר זיהוי. מידע מזהה (identifying data) כולל פרטים מזהים ישירים, למשל: שם, מספר טלפון, כתובת IP של מחשב, צילום פנים, הקלטת קול וכדומה. מידע שמאפשר זיהוי עלול להתרחש כאשר מתאפשרת החזרת זיהוי (re-identification), כך שמאמץ סביר מאפשר לזהות את המשתתפים גם בהיעדר מידע מזהה ישיר, למשל: כשמידע משולב ממספר מקורות (כגון מידע מסקר ורשומות מנהליות), או כשמידע שמקורו במדגמים קטנים כולל מידע איכותני עשיר (למשל: תשובות לשאלות פתוחות, כפי שניתנו בסקר).

מחד, היכולת לשלב מספר מקורות מידע מעשירה את הידע והתובנות שניתן להפיק מנתוני סקר, לדוגמה: מידע מנהלי (כגון גובה הכנסות) ומידע הקשרי (מידע קונטקסטואלי כגון מיקום גיאוגרפי של

²⁷ United Nations, Economic Commission for Europe Conference of European Statisticians (2024). Follow-up on the Conference of European Statisticians seminar "Data ethics – a key enabler of social acceptability." Seventy-second plenary session, Geneva. 20–21 June 2024. ECE/CES/2024/19. Retrieved from https://unece.org/sites/default/files/2024-06/ECE_CES_2024_19_E.PDF

המשיב, נתוני חיישנים מהטלפון החכם). מאידך, שילוב של מקורות מידע שונים על אותו הנדגם מגביר את הסיכוי לזיהויו ובכך מגביר את הסיכון לפגיעה בפרטיות.

הסכמה מדעת

בתקנות האירופאיות לשמירה על מידע אישי (Regulation (EU), 2016 – GDPR)²⁸ מצוין כי טיפול במידע מזהה, אם אינו נעשה מכוח החוק, צריך להיות על בסיס הסכמה מדעת (שם, עמודים 6–8 בעיקר), מתוך בחירה חופשית של הנדגם ותוך הבנה ברורה וחד-משמעית של מי שנותן את הסכמתו כי התהליך הסטטיסטי כולל מידע מזהה על אודות הנדגם. הבקשה להסכמה מדעת יכולה להיעשות באמצעים שונים (בכתב, באופן אלקטרוני, בעל פה) ועליה לכסות את כל הפעולות הסטטיסטיות שאמורות להתבצע במסגרת המטרות שנקבעו לפעולה הסטטיסטית.

האחראי לפעולה הסטטיסטית צריך להיות מסוגל להוכיח שהמשתתף (נדגם או מרואיין חליפי) נתן הסכמה מדעת, להבטיח שהמשתתף מודע למתן ההסכמה וכי הוא יודע למה ההסכמה ניתנה. מסמך הבקשה להסכמה מדעת אמור להיות מוצג בנפרד מאמצעים אחרים (כגון משאלון הסקר), מנוסח בצורה בהירה וקלה להבנה ובשפה ברורה ופשוטה ולא לכלול תנאים לא-הוגנים. כדי שהסכמה תיחשב להסכמה מדעת, על המשתתף לקבל מידע, לכל הפחות על אודות האחראי לפעולה הסטטיסטית ועל מטרות השימוש במידע המזהה. הסכמה חופשית כוללת אפשרות לבחור אם להשתתף או לסרב, או להשתתף ולאחר מכן להפסיק השתתפות, כל זאת בלי חשש ובלי להיפגע. כמו כן, אסור לאחראי לפעולה הסטטיסטית להתנות מתן שירות לנדגם בכך שהוא יסכים להשתתף.

משתתף רשאי להורות למחוק מידע מזהה אודותיו אם המידע כבר אינו נחוץ להשגת המטרה המוצהרת של הפעולה הסטטיסטית או במקרה שהמשתתף נסוג מההסכמה שנתן (שם, עמוד 12). בכל מקרה, יש לאפשר לו לסגת מהסכמה בכל עת ובאותה קלות כמו מתן הסכמה (שם, עמוד 37). לבסוף, יש לציין כי אם הנדגמים הם קטינים או חסרי ישע, הסכמה מדעת תתקבל רק על ידי אפוטרופוס טבעי (למשל: הורה), או מי שמונה על ידי בית משפט (שם, עמוד 37).

בנוסף לתקנות האירופאיות לשמירה על מידע אישי, בבריטניה פורסמו הנחיות אתיקה לביצוע מחקרים בתחום החברה וההתנהגות על ידי גופי ממשל²⁹. בהנחיות אלה ניתן למצוא פירוט רב על

²⁸ Regulation (EU) (2016) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), *Official Journal of the European Union* L, 119(1), pp. 1–88. Retrieved from <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679>.

²⁹ Government Social Research. (2021). *GSR Professional Guidance: Ethical Assurance for Social and Behavioural Research in Government*. Retrieved from <https://www.gov.uk/government/publications/ethical-assurance-guidance-for-social-research-in-government>.

היבטים שונים שקשורים להתנהלות אתית במחקר, לרבות הסכמה מדעת. כפי שנאמר לעיל, הסכמה מדעת של הנדגם נדרשת בכל מקרה שבו אין חובה חוקית של נדגם (אדם או ארגון) למסור מידע אישי. הסכמה מדעת צריכה להינתן מראש, לאחר שהנדגם קיבל מידע ברור על מטרת הסקר, על מבצע הסקר ומזמינו, ועל השימושים שיעשו במידע שנאסף. לנדגם צריכה להיות אפשרות לחשוב ולשאול שאלות, בטרם ייתן את הסכמתו. גם לאחר מתן ההסכמה, הנדגם יכול לחזור בו בכל שלב בסקר. הפנייה לקבלת הסכמה מדעת צריכה להיעשות בנפרד מהסקר ולהיות מתועדת בנפרד ובאופן מאובטח. הסכמה מדעת צריכה להיות ספציפית לפעולה, למשל: אם היא ניתנה עבור גל איסוף אחד, נדרש לבקש הסכמה מדעת נוספת עבור גל איסוף אחר ו/או עבור השתתפות בראיון חוזר ו/או בסקר אחר, גם אם הדבר נעשה על ידי אותו מבצע ועבור אותו מזמין. בדומה, שימוש משני בנתונים צריך להיעשות בכפוף להסכמה מדעת ולמידע שהמשתתף קיבל באותו זמן שהוא נתן את הסכמתו. כל שינוי שנעשה במידע מחייב בקשה נוספת להסכמה מדעת של המשתתף.

הסכמה מדעת באמצעות מרואיין חליפי תיעשה אם הנדגם אינו מסוגל להבין את ההשלכות של השתתפותו בסקר ואינו יכול לתת הסכמה מדעת וכן במקרים שבהם מדובר בקטינים.

מכשולים והטיות שקשורים בהסכמה מדעת³⁰

השימוש במסמך הסכמה מדעת צריך להיעשות תוך תשומת לב למספר היבטים. ראשית, על מנת שהמשתתף יוכל לשקול אם להסכים לבקשה, מסמך הבקשה להסכמה מדעת אמור להיות מנוסח באופן ברור וקל להבנה ולכלול את כל המידע הרלוונטי על הפעולה הסטטיסטית (למשל: מטרת הפעולה, קישור למקורות מידע נוספים, גורמים שמעורבים בטיפול במידע ושימושים שיעשו במידע); כן מצופה כי המשתתף יקרא את כל הכתוב ויבין את משמעות הדברים. בפועל, מסמך הבקשה להסכמה מדעת עלול לכלול מונחים שאינם ברורים לקורא, להיות ארוך ולא לכלול את כל השימושים האפשריים במידע, במיוחד כשמדובר בנתוני עֵתֶק (big-data), שאופן עיבודם וניתוחם ממשיך להתפתח ולגלות אפשרויות חדשות. הקורא עלול שלא להבין חלקים מהמסמך או אפילו לדלג על חלקים בו, או על כולו. לכן, בעיקר כשהמסמך נקרא על ידי המשתתף ללא תיווך מראיין, קשה לדעת אם הסכמה אכן ניתנה מדעת ואם היא עומדת בדרישות החוק והאתיקה.

שנית, מסמך הבקשה להסכמה עלול לגרום להטיות דגימה, עקב אינטראקציה עם מאפייני המשיב, למשל: אם בעלי השכלה נמוכה עלולים לגלות הבנה פחות טובה לכתוב, הרי שעלולה להתקבל

³⁰ להרחבה ראו:

Plutzer, E. (2019). Privacy, sensitive questions, and informed consent: Their impacts on total survey error, and the future of survey research. *Public Opinion Quarterly*, 83(SI), 169-184.

Struminskaya, B., & Sakshaug, J. W. (2023). Ethical considerations for augmenting surveys with auxiliary data sources. *Public Opinion Quarterly*, 87(SI), 619-633.

התפלגות שונה של מסכימים ולא מסכימים ברמות השכלה שונות. דוגמה נוספת: עמדות של המשתתף בנוגע לפרטיות באופן כללי, או לתחומים פרטניים בכתב ההסכמה (למשל: שיתוף מידע מחיישנים או מאפליקציות בטלפון חכם) ומידת אמון שהוא חש כלפי האחראי לפעולה הסטטיסטית, כל אלה עלולים להשפיע על החלטתו אם להסכים, או לסרב, לבקשה להסכמה מדעת.

שלישית, כאשר התכנים הנבדקים נתפסים כרגישים או כמפירים נורמות חברתיות (למשל: נטייה מינית, צריכת סמים ועוד), ה"מסכימים", מתוך רצון לשמור על דימוי חיובי, עלולים לספק תשובות "נורמטיביות" ופחות כנות, לעומת מצב שבו היו משיבים באופן אנונימי, או לעומת המסרבים (לו היו משיבים).

מאמצים לצמצום הפגיעה בפרטיות ולשמירה על סודיות הנתונים³¹

זיהוי של נדגמים מתרחש כאשר נאסף מידע על משתנים מזהים ישירים (כגון שם, מספר תעודת זהות וכדומה) אך עלול להתרחש גם כשמדובר במזהים עקיפים (אשר מכונים גם "מזהים למחצה", כגון גיל, השכלה, כתובת מגורים, כתובת סניף קופת החולים, מידע על מצב בריאות ועוד). מזהים עקיפים עצמם אינם מזהים את הנדגם, אך שילוב של מספר משתנים כאלה ו/או הצלבתם עם מידע ממקור אחר מגבירים את היכולת להחזרת זיהוי של נדגם.

גופי ממשל, כגון הלשכה המרכזית לסטטיסטיקה ומשרד הבריאות, אשר משתמשים במידע מזהה מתוקף תפקידם, עושים מאמץ רב לשמור על פרטיות הנדגמים ועל סודיות הנתונים, ובד בבד פועלים לשיתוף הנתונים לצורכי מחקר ולתועלת הציבור. שיתוף הנתונים נעשה לאחר עילום (או: "התממה", anonymization), שהוא תהליך שכולל הפיכתם של נתונים אישיים מזהים לנתונים שאינם מזהים, באמצעים שונים. עילום הנתונים נעשה תוך איזון בין הצורך להגן על פרטיות הנדגמים לבין התועלת שניתן להפיק מהנתונים ותוך התחשבות במגבלות שונות, למשל: אופי הנתונים (למשל: אלו משתנים מזהים ישירים ומשתנים עקיפים יש בקובץ) והשימושים המיועדים במידע (למשל: למחקר או לשימוש הציבור). ככל שמטרות הפעולה הסטטיסטית דורשות מידע פרטני על הנבדקים, העילום ייעשה תוך שמירת הנתונים ברמת הנדגם (individual level data) באמצעות הסרת זיהוי (de-identification), למשל: הסרה, הסתרה או החלפת משתנים שמאפשרים זיהוי, החלפת מידע בין שדות, הוספת

31 להרחבה ראו:

בר־זיו, ש' וז'רסקי, ט' (2019). פרטיות במשבר זהות: אסטרטגיות הסדרה בעידן ההתממה. משפט, חברה ותרבות, ב, 125–166. אוצר מתוך:

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3350266

משרד הבריאות (2018). מסקנות הועדה ליישום המלצות השימושים המשניים במידע https://www.health.gov.il/PublicationsFiles/health_info.pdf בריאות. אוצר מתוך:

Gadotti, A., Rocher, L., Houssiau, F., Crețu, A. M., & de Montjoye, Y. A. (2024). Anonymization: The imperfect science of using data while preserving privacy. *Science Advances*, 10(29), eadn7053.

"רעש" ועוד. מנגד, כשיש צורך גבוה בשמירה על פרטיות הנדגמים, נעשה מאמץ למניעה של החזרת זיהוי (re-identification), דבר שדורש הליכים של שמירת המידע ברמת הקבוצה (כגון שימוש בממוצע, חציון, ערך מתאם ועוד) או על ידי יצירת קובץ סינתטי, דבר שגורם לצמצום עושר המידע שניתן להפיק מהנתונים.

מומלץ כי שמירת מידע שמאפשר זיהוי תיעשה בהתאם לרמת הסיכון לפגיעה בפרטיות הנדגמים וסודיות המידע. ניתן לאפיין רמות סיכון, בהתאם לגורמים שונים (למשל: סוג המשתנים, מצב עילום הנתונים, גודל המדגם בתת-קבוצה):

סיכון ודאי או גבוה – מידע פרטני (individual level data) הכולל זיהוי אישי ישיר (כגון שם ומספר זיהוי), או מידע פרטני הכולל זיהוי למחצה (כגון גיל וכתובת) שמאפשר החזרת זיהוי באמצעות שילוב מידע או על ידי הצלבתו עם מידע מקובץ אחר;

סיכון סביר – מידע פרטני, ללא זיהוי אישי ישיר אך עם זיהוי למחצה, שעבר עילום (anonymization) ברמת הנדגם (למשל: החלפה או הסרת משתנים) כך שקיים סיכון לא מבוטל להחזרת זיהוי;

סיכון נמוך – מידע שעבר עילום מעבר לרמת הנדגם, כגון מידע מקובץ או מידע סינתטי, אך אופי המידע מגביר סיכון להחזרת זיהוי (למשל: תכנים רגישים, מעט נדגמים, ניתוח נתונים בסביבה אינטראקטיבית);

סיכון נמוך ביותר – מידע שעבר עילום מעבר לרמת הנדגם, כגון מידע מקובץ או מידע סינתטי, ואופי המידע מקטין סיכון להחזרת זיהוי (למשל: תכנים לא רגישים, ריבוי נדגמים, ניתוח נתונים בסביבה לא-אינטראקטיבית).

בכל מקרה שבו נתונים עברו עילום, יש צורך במאמץ פעיל ומתמשך לוודא כי נתונים אלה עמידים בפני זיהוי מחדש. בנוסף, יש צורך באבטחת מידע לנתונים, בהתאם לרמת הסיכון שקיימת לזיהוי של הנדגמים. אבטחת מידע כוללת התייחסות להיבטים שונים, למשל: אבטחה פיזית וטכנית של המחשבים ושל המידע ששמור בהם; שימוש בכלים טכנולוגיים להתמודדות עם נזקות ועם ניסיונות פריצה; שימוש במערכות תיעוד, פיקוח ובקרה.

ניתן לנסות לצמצם את הפגיעה בפרטיות הנדגמים, כבר בשלב איסוף המידע, למשל: בסקרים שנעשים בטלפון חכם, ניתן להשתמש במידע מקובץ, במקום במידע פרטני (כגון מידע בינרי על קיום, או אי-קיום, של שיחות בטלפון נייד, במקום הקלטת תוכן השיחות; המרת מיקומים גיאוגרפיים למרחקים וכדומה). בנוסף, ניתן לאפשר למשתתפים לצפות במידע המקובץ, לפני שהם מאשרים את שליחתו לאחראים לסקר.

נטל השבה הוא המידה שבה המשיב תופס את השתתפותו בסקר כקשה, גוזלת זמן או מעיקה מבחינה נפשית (Graf, 2008). ניתן לאפיין שני סוגים של נטל השבה: ממשי ונתפס. נטל השבה ממשי נוצר מגורמים אובייקטיביים, כגון אורך הראיון, תדירותו (בסקרים חוזרים) והמורכבות הקוגניטיבית שנדרשת מהמשיב (למשל: שאלות שדורשות שלילת מידע מהזיכרון, בחירה בין אפשרויות דומות או דירוגן לפי חשיבות וכדומה). נטל השבה נתפס נוצר מגורמים סובייקטיביים, כגון מוטיבציה של הנדגם לשתף פעולה, תחושתו ביחס לקושי, לזמן ולמאמץ שנדרש ממנו ותחושת לחץ כתוצאה משאלות שנתפסות על ידו כפולשניות (Bottone Modugno, & Neri, 2021).

מדדים לנטל ממשי כוללים, למשל: מספר עמודי השאלון, מספר השדות בשאלון, משך הזמן שנדרש להשיב על הסקר, ובסקרים ארגוניים גם הזמן שנדרש לארגון של המידע, מספר האנשים שתרמו להשלמת השאלון והאם נדרש סיוע מגורמי חוץ להשלמת המטלה. מדדים לנטל נתפס כוללים, למשל: תפיסות ועמדות של המשיבים, כגון מידת העניין או החשיבות שהמשיבים מייחסים לסקר, מידת האמון שהם חשים כלפי מזמין הסקר ו/או עורך הסקר, הערכתם את הזמן שנדרש להשלמת הסקר כארוך או קצר ומידת הקושי, או המורכבות, שהם חשו בעת המענה. נוסף לכך, יש חוקרים שמסיקים על נטל ההשבה מתוך תוצאות לא רצויות בסקר, כגון הפסקת מענה במהלך הסקר וסירוב להשתתף בסקר חוזר (Yan et al., 2020).

- Bottone M., Modugno, L., & Neri, A. (2021). Response Burden and Data Quality in Business Surveys. *Journal of Official Statistics*, 37(4), 811-836.
- Dale, T., & Haraldsen, G. (2007). *Handbook for Monitoring and Evaluating Business Survey Response Burdens*. European Commission, Eurostat
- Graf, I. (2008). Respondent burden. In P. Lavrakas (Ed.), *Encyclopedia of survey research methods* (p. 739). Thousand Oaks, CA: Sage Publications.
- Yan et al. (2020). Response burden: What is it and what predicts it?. *Advances in Questionnaire Design, Development, Evaluation and Testing*, 193-212

יסודות להנחיות אתיקה

בישראל, בדומה למדינות רבות, אין תקנות או הנחיות אתיקה אחידות שמיועדות לסקרים. עם זאת, מעין במסמכים שפורסמו על ידי גופים לאומיים ומחקריים בארץ ובעולם, ניתן למצוא דמיון בין ערכים, עקרונות וכללים שנוסחו לשמירה על אתיקה.

בעמודים הבאים יובאו תמציות של הנחיות שפורסמו על ידי גופים סטטיסטיים וגופי מחקר ואקדמיה.

גופים סטטיסטיים

הלשכה המרכזית לסטטיסטיקה והאיגוד הישראלי לסטטיסטיקה ומדע הנתונים מפרטים את כללי האתיקה הללו:³³

1. חובות כלליות – ביצוע עבודה באופן מקצועי, ביושר, תוך הישענות על עקרונות מדעיים; קידום אמון הציבור בעשייה הסטטיסטית; פרסום תרומתה של העשייה הסטטיסטית; שקיפות של פעולות סטטיסטיות שנעשו ומידע שקשור בהן.
2. חובות כלפי החברה – הימנעות מפעולות סטטיסטיות שעלולות לגרום להטיה והטעיה; פרסום ממצאים באופן מקיף ומהימן; מסירת ממצאים סטטיסטיים לציבור בליווי הסברים כדי לוודא שהציבור יעריך נכון את מהימנותם ויבין אם מדובר בממצאים ישירים, בפרשנות או בתחזיות.
3. חובות כלפי המשיב ו/או מי שעליו נאסף המידע – איסוף מידע רלוונטי בלבד, תוך צמצום פגיעה בצנעת הפרט; מסירת מידע רלוונטי למשיב על הפעולה הסטטיסטית; יידוע על חובת השתתפות ובהיעדרה – קבלת הסכמה מרצון; שמירה בהתאם לדין על סודיות המידע שנאסף.
4. חובות כלפי המזמין – הגדרה מראש של זכויות הצדדים וחובותיהם; סיכום עם המזמין על אופן פרסום הממצאים; איסור התניית תוצאות (באמצעות פעולות שתוצאותיהן הוכתבו מראש או שיש בהן הטיה או הטעיה מכוונת); תיאור שיטות לביצוע הפעולות הסטטיסטיות והשלכותיהן; גילוי נאות על המבצעים שהינו רלוונטי לביצוע הפעולות הסטטיסטיות.
5. חובות כלפי עמיתים למקצוע – שקיפות של שיטות העבודה לשם ביקורת מקצועית; מתן אפשרות לניתוח משני ללא פגיעה בזכויות של מי שהמידע נאסף אודותיו; הכללת עקרונות אתיים אלה על כל השותפים לביצוע הפעולה הסטטיסטית והתחשבות בכללי אתיקה של בעלי מקצוע אחרים השותפים לפעולה.

³³ הלשכה המרכזית לסטטיסטיקה (אין תאריך). כללי האתיקה של הלשכה המרכזית לסטטיסטיקה. אוצר מתוך https://www.cbs.gov.il/he/About/Documents/ethics_h.pdf

המועצה הציבורית לסטטיסטיקה – האיגוד הישראלי לסטטיסטיקה (אין תאריך). כללי אתיקה מקצועית לעוסקים בסטטיסטיקה. אוצר מתוך <https://statistics.org.il/about>

הנחיות, הדומות לאלה המתוארות לעיל, פורסמו על ידי האיגוד האמריקאי לסטטיסטיקה

(ASA - American Statistical Association, 2022):³⁴

1. יושרה מקצועית ואחריות – שיקול דעת לגבי יכולת ביצוע והשלמת המשימות; שימוש בשיטות ובנתונים תקפים, רלוונטיים ומתאימים; שימוש בפעולות סטטיסטיות שלא יפגעו באוכלוסיות חלשות; התנגדות להתניית תוצאות; קבלת אחריות מקצועית; מתן הכרה מקצועית לשותפים; שקיפות לגבי ניגוד עניינים; יחס של כבוד ושוויון כלפי כל האנשים; תגובה לחריגה של אחרים מהנחיות האתיקה; שמירה על כשירות מקצועית; התנהלות בהתאם למדיניות, תקנות וחוקים; שימור, כיבוד וקידום הנחיות האתיקה.
2. שימוש ראוי בנתונים ושיטות – שקיפות לגבי מקורות הנתונים והשימוש בהם; שקיפות ביחס להנחות שקשורות להתנהלות הסטטיסטית, כגון שיטות, מגבלות, מקורות לטעות והטיות; דיווח על המטרות והשימושים הסטטיסטיים המתוכננים והסופיים והתאמות שנעשו; מתן אפשרות לשימוש משני בנתונים, ללא פגיעה בזכויות של נשואי המידע; תיקון של טעויות משמעותיות, גם לאחר פרסום הממצאים; בדיקות תוקף חוזרות לפעולות סטטיסטיות שנעשות לאורך זמן.
3. אחריות כלפי בעלי עניין – הגדרת היעדים תוך התחשבות בצרכים של בעלי עניין, כדי להשתמש בפעולות סטטיסטיות רלוונטיות; שימוש בפעולות סטטיסטיות שלא יגרמו להטעה או להטיה; מסירת ממצאים באופן ברור כדי שיובנו כראוי ויאפשרו את יישומם; יידוע לגבי מגבלות אפשריות של שימוש בפעולות הסטטיסטיות תוך התייחסות לגורמים שעלולים להשפיע על יעילותן; הסבר להשלכות שליליות אפשריות במקרה של סטייה משיטות שהוסכם עליהן; שאיפה להנגיש ידע מתודולוגי חדש לתועלת הציבור; הבנה וציות לדרישות הסודיות הנוגעות לאיסוף, מסירה והפצה של מידע ולכל מגבלה על שימוש במידע.
4. אחריות כלפי משתתפים, נשואי המידע שנאסף וכל מי שמושפע ישירות מהפעולות הסטטיסטיות – שמירה על חוקים, תקנות והנחיות לשמירה על המשתתפים; הימנעות מנטל וסיכון משיבים תוך שימוש במדגמים לא גדולים מידי ואיסוף מידע הכרחי בלבד וכן צמצום גודלי מדגמים וכמות הניסויים בניסויים על בעלי-חיים; שמירה על זכויות ופרטיות נשואי המידע ועל סודיות המידע בהתאם לדין; שימוש במידע תוך שמירת זכויות המשתתפים וכפי שהתירו (אם נדרשה הסכמה מדעת), לרבות שימוש ראשוני ושניוני במידע; שימוש מועיל ומיטבי במידע שנאסף; הכרה במגבלות החוק לשמירה על פרטיות וסודיות.
5. אחריות כלפי עמיתים בצוותים רב-תחומיים – כיבוד מחויבויות ותקני אתיקה של מומחים מתחומים אחרים; מתן עדיפות להנחיות ה-ASA במקרה של התנגשות עם הנחיות אחרות;

³⁴ American Statistical Association (2022). *Ethical Guidelines for Statistical Practice*. Retrieved from <https://www.amstat.org/your-career/ethical-guidelines-for-statistical-practice>.

שקיפות לגבי כל הפעולות הסטטיסטיות שנעשו; הימנעות מפשרות שיפגעו בתוקף וזהירות משימוש בפעולות סטטיסטיות לא ראויות.

6. אחריות כלפי עמיתים וכלפי המקצוע – התחשבות במומחיות וניסיון של סטטיסטיקאים אחרים; הכרה בתרומה של עמיתים; מתן סיוע, ייעוץ ומשוב לאחרים; לקיחת אחריות על ייעוץ, הנחיה והדרכה לביצוע פעולות סטטיסטיות; מתן אפשרות לשחזור פעולות וממצאים דרך שיתוף נתונים, שיטות ותיעוד, במידת האפשר; לשמש נציג לשיטות ופעולות סטטיסטיות ראויות.

7. אחריות של סטטיסטיקאים מובילים – וידוא שימוש ראוי בפעולות סטטיסטיות, בהתאם להנחיות ה-ASA; קידום אווירת עבודה מכבדת ומועילה; עידוד לשיתוף פעולה, התפתחות מקצועית והכרה בתרומה של סטטיסטיקאים; יצירת תרבות של תיקוף הנחות והערכת מודלים.

8. אפשרות של התנהלות לא ראויה – מודעות להיבטים שונים של התנהלות לא-ראויה, הימנעות מכך ועידוד אחרים להימנע מכך; הימנעות מהאשמת אחרים ללא בסיס; ניהול הליך שקוף והוגן לבחינת התנהלות לא-ראויה; הימנעות מדיון פומבי בטרם הסתיימו הליכי הבירור; סיוע לחזור לשגרה בהתאם לתוצאת ההליכים; תמיכה במי שאחראי לחשיפת התנהלות לא-ראויה.

הנחיות אתיקה פורסמו לאחרונה על ידי הרשות לסטטיסטיקה של קנדה (Beaulieu, 2024)³⁵ והן כוללות היבטים אלה:

1. טובת הציבור הקנדי – הגדרה והדגמה של ההטבות המצופות, כולל דוגמאות מוחשיות; אם הפסקת הפרויקט תלווה בתוצאות לוואי שליליות, יש להסביר ולתאר את השלכותיהן, כדי להצדיק את הפרויקט.
2. פרטיות וסודיות – נקיטת צעדים כדי להימנע מפריצה, או מהשפעות אפשריות של פריצה, למידע; איסוף של מידע הכרחי בלבד.
3. שקיפות ואחריות – ייזום פעולות לקידום השקיפות; אחריות בעת שימוש בלמידת מכונה (ML) או באינטליגנציה מלאכותית (AI).
4. אמון וקיימות – שמירה על קשרים עם שותפים ועם ספקי מידע; דאגה לכך שכללים אלה יקוימו.
5. איכות הנתונים – הכרת מקורות לטעות ו/או הטיה ודרכים להתמודד איתם; ההטבות המצופות מהפרויקט תלויות באיכות הנתונים.

³⁵ Beaulieu, M. (2024, March). *The Role of Data Ethics to Maintain and Improve Public Trust: The Statistics Canada Experience*. Workshop on Ethics in Modern Statistical Organisations, Geneva, Switzerland. Retrieved from https://unece.org/sites/default/files/2024_04/Ethics2024_S3_Canada.

6. הוגנות והימנעות מנזק – הימנעות מיצירת סטראוטיפים (בעיקר כשהמידע נאסף על קבוצות אוכלוסייה פגיעות או בשוליים); שימוש במידע שנאסף מצדיק את אחזקתו (כולל תכניות לשימוש משני במידע למטרות שונות).

בבריטניה נכתבו מסמכי אתיקה על ידי גופים מרכזיים שעוסקים בסטטיסטיקה לרבות הרשות הבריטית לסטטיסטיקה (UK Statistics Authority), החברה המלכותית לסטטיסטיקה (Royal Statistical Society) והמשרד המרכזי לדיגיטציה ולמידע (Central Digital and Data Office). במסמך של הרשות הבריטית לסטטיסטיקה (UK Statistics Authority, n.d.)³⁶ נכללים עקרונות דומים לאלה שנכללים בהנחיות הקנדיות:

1. טובת הציבור – לשימוש במידע יש תועלת ברורה עבור המשתמשים והדבר משרת את טובת הציבור.
 2. פרטיות וסודיות – פרטיות נשואי המידע שנאסף (אנשים או ארגונים) נשמרת, המידע נשמר באופן חסוי ומאובטח, השימוש בהסכמה מדעת נעשה בהתאם לצורך.
 3. שיטות ואיכות – הסיכונים והמגבלות של טכנולוגיות חדשות נשקלים וקיים די הצורך פיקוח אנושי, כך שהשיטות שמופעלות מתאימות לסטנדרטים מקובלים של יושרה ושל איכות.
 4. ציות לחוקים – השימוש במידע והשיטות המופעלות הולמים את דרישות החוק הקיימות.
 5. מעורבות ציבורית – התחשבות בדעות הציבור ביחס למידע שנעשה בו שימוש ולתועלות של המחקר.
 6. שקיפות – הנגישות למידע, השימוש בו ושיתופו נעשים באופן שקוף ומדווחים באופן ברור ונגיש לציבור.
- היבטים משלימים לאלה שתוארו לעיל, ניתן למצוא בקוד האתי של החברה המלכותית לסטטיסטיקה (Royal Statistical Society, 2014)³⁷, אשר ממוקד, רובו ככולו בהתנהלות הסטטיסטיקן, להלן:

א. טובת הציבור –

1. בראש ובראשונה להיות אחראי כלפי הציבור, לרבות בריאותו, בטחונו וסביבתו.

2. לשמור על זכויות אדם בסיסיות ולהימנע ממעשים שפוגעים בזכויות אלה.

ב. חובות כלפי מעסיקים ולקוחות –

3. לבצע עבודה בתשומת לב ובזהירות בהתאם לדרישות המעסיק או הלקוח.

³⁶ UK Statistics Authority (n.d.). *Ethical Principles*. Retrieved from <https://uksa.statisticsauthority.gov.uk/the-authority-board/committees/national-statisticians-advisory-committees-and-panels/national-statisticians-data-ethics-advisory-committee/ethical-principles>.

³⁷ Royal Statistical Society (2014). Code of Conduct, RSS. Retrieved from <https://rss.org.uk/RSS/media/File-library/About/2019/RSS-Code-of-Conduct-2014.pdf>

4. לכבד כל הסכם לשמירה על סודיות עם מעסיק או לקוח.
 5. לסרב להיות חתום על עבודה שלא נעשתה על ידו של הסטטיסטיקן או שתרומתו מוצגת בצורה מטעה.
 - ג. חובות למקצוע ולאגוד –
 6. לתמוך בשמם הטוב של המקצוע ושל האגוד.
 7. לחפש דרכים לקדם את הידע וההבנה של מדע הסטטיסטיקה ושל השימושים שבו.
 8. לשמור על יושרה כלפי עמיתים וכלפי חברים ממקצועות אחרים שמשותפים איתם פעולה.
 9. לקיחת אחריות אישית על עבודה שנושאת את שמו של הסטטיסטיקן.
 - ד. מומחיות ויושרה –
 10. לשאוף לפעול באופן הגון וביושרה בכל תחומי הפעילות המקצועיים.
 11. לשמור על פיתוח מקצועי מתמשך בהתאמה לקוד האתי.
 12. לשאוף לפעול בהתאם להנחיות מוכרות לביצוע מיטבי.
 13. לדווח לחברה המלכותית לסטטיסטיקה (Royal Statistical Society) על כל הרשעה פלילית של הסטטיסטיקן לרבות אלימות, חוסר יושר או התנהלות מקצועית לא-הולמת.
- הנחיות של האיגוד הבינלאומי לסטטיסטיקה³⁸ מיועדות לכל מי שעוסק בסטטיסטיקה, במטרה לסייע לקבל החלטות אתיות מתאימות. ההנחיות מאורגנות בשני חלקים: ערכים משותפים ועקרונות אתיים.
- א. ערכים משותפים לעוסקים בסטטיסטיקה:
1. כבוד – כיבוד פרטיות האחר והבטחת סודיות המידע; כיבוד הקהילות שמהן נאסף המידע וזהירות מפגיעה בהן דרך שימוש לרעה בממצאים; כיבוד עבודתם של אחרים.
 2. מקצועיות – הבנת צורכי המשתמש ופיתוח פתרונות מתאימים; שימוש בידע ובכלים מקצועיים לשם שירות הציבור; שמירה על איכות גבוהה של העבודה הסטטיסטית; התאמת המידע והשיטות למטרות הפעולה הסטטיסטית; דאגה להשתמש במידע בהתאם למטרות שהוגדרו ולנורמות חברתיות; שמירה על אובייקטיביות; ציות לחוק וקידום שינוי של חוקים שעלולים לפגוע בעבודה סטטיסטית טובה; שמירה עדכון מקצועי מתמשך; פיתוח שיטות חדשות; זהירות מפעולה עם ניגוד עניינים; עבודה אחראית עם מעסיקים.
 3. אמון ויושרה – עבודה ללא לחץ מפוליטיקאים או מממנים; שקיפות לגבי השיטות הסטטיסטיות שמופעלות; שאיפה להפיק תוצאות באופן אובייקטיבי; הצגת מידע ושיטות באופן ישר ופתוח, לרבות מגבלות הממצאים; לקיחת אחריות על פעולות; כבוד לקניין

³⁸ International Statistical Institute (2023). *Declaration on Professional Ethics*, ISI. Retrieved from <https://isi-web.org/declaration-professional-ethics>

רוחני; חתירה לרעיונות חדשים ולדיוק אמפירי; כיבוד אמות מידה אובייקטיביות ומבוססות להערכה.

ב. עקרונות אתיים:

1. חתירה לאובייקטיביות – בחירה ושימוש בשיטות המתאימות ביותר, ללא חשש או שיקולים זרים ותוך התחשבות בגורמים כגון דיוק ועמידה בלוח זמנים; וידוא שהמידע שמנותח מתאים למטרות שהוגדרו ושלסטטיסטיקן יש את הידע הנדרש בתחום התוכן; הצגת ממצאים בצורה פתוחה, שלמה ושקופה; מתן מידע לגבי מגבלות הממצאים; זהירות מפני פרשנות לא נכונה של הממצאים או שימוש לרעה בהם; פרסום הממצאים לטובת הציבור הרחב וזהירות שלא לפגוע בקבוצות בתוכן.
2. הגדרת תחומי אחריות – יש לברר, להבין ולפרט אלו היבטים (שקשורים לתפקידים ולמחויבויות של המעסיק, הלקוח או המממן ושל הסטטיסטיקן) עלולים לעורר קשיים אתיים; על הסטטיסטיקן לפעול בתוך תחומי מומחיותו ובהתאם לצורך ולאפשרות לבקש סיוע ממומחים אחרים.
3. שקילות חלופות באופן אובייקטיבי – שקילה של שיטות והליכים אפשריים, יידוע של המעסיק, הלקוח או המממן לגביהם ומתן המלצה על החלופה הנבחרת.
4. ניגוד עניינים – יש להימנע מלקיחת משימת שתוצאותיהן יובילו לניגוד ענייניים פיננסי או אישי.
5. הימנעות מהתניית תוצאות – יש לדחות כל ניסיון להתנות תוצאות מראש, לרבות התניית תשלום בתוצאות.
6. שמירה על סודיות – יש לשמור על סודיות של מידע חסוי (privileged).
7. הפגנת כישורים מקצועיים – על הסטטיסטיקן לשאוף לעדכן את ידיעותיו וכישוריו המקצועיים, תוך תשומת לב להתפתחויות של טכנולוגיות, הליכים ותקנים רלוונטיים ולעודד אחרים לנהוג כך.
8. קידום אמון הציבור בסטטיסטיקה – דאגה לתיאור מדויק ונכון של התוצאות וזהירות בפרשנות הממצאים, תוך הפניית תשומת-לב של המשתמשים למגבלות אפשריות של מהימנותם ויישומם, בעזרת אמצעים כמותיים ואיכותניים.
9. הצגה של שיטות ושל ממצאים והערכתם – יש להפיץ לציבור מידע ראוי, על מנת שניתן יהיה להעריך את השיטות, ההליכים והממצאים בצורה בלתי-תלויה.
10. פרסום עקרונות אתיים – יש לוודא שהעקרונות האתיים יובנו כראוי על ידי כל מי שקשור לפעולה הסטטיסטית.
11. לקיחת אחריות על היושרה של המקצוע – יש לציית לחוקי המוסר שכרוכים בהתנהלות מדעית ואקדמית, לרבות הימנעות מרמייה, שיבוש או הסתרת התנהלות לא ראויה ומחסימת פעולה מחקרית של אחרים.

12. הגנה על נבדקים – שמירה על נבדקים מפני השפעות מזיקות אפשריות, כך שבאיסוף מידע רגיש, יש לוודא שקיימת הצדקה לפעולה, לקבל הסכמה מדעת ולשמור על סודיות המידע. עם זאת, הסכמה מדעת אינה פוטרת מאחריות זו.

גופי מחקר ואקדמיה

היבטים משלימים, ובחלקם חופפים, פורסמו על ידי האיגוד האמריקאי לחקר דעות הציבור (American Association for Public Opinion Research, 2021)³⁹, להלן תמצית הדברים:

א. עקרונות של אחריות מקצועית:

1. אחריות כלפי המשתתפים – הימנעות משיטות שמזיקות למשתתפים, שמסכנות, שמשפילות, או שמטעות אותם; היעדר מטרות חבויות (כגון גיוס כספים, פעילות פוליטית וכד') במחקר; השתתפות במחקר היא מבחירה, למעט במקרה שהחוק דורש השתתפות; מידע לא מטעה על מזמין המחקר ומטרותיו, למעט במקרים שהדבר עלול לפגוע במשתתפים או להטות את תשובותיהם; שמירה על זכויותיהם של מיעוטים וקבוצות מוחלשות; התנהלות בהתאם לחוקים ולתקנות.
2. אחריות בעת איסוף מידע אישי מזהה – יידוע של מוסרי המידע על אודות השימוש שייעשה בנתונים המזהים; טיפול בנתונים בהתאם להנחיות, לחוקים ולתקנות; מניעת גישה לנתונים מזהים והשמדתם בתום המחקר ובהתאם לחוק; הסתרת כל מידע שיכול לזהות משתתפים, למעט במקרה שהם נתנו את רשותם לכך; אישור של מוסרי המידע לשימוש במידע מזהה למטרות אחרות; קבלת אחריות לשמירת פרטיות המשתתפים וסודיות המידע, למעט אם המשתתפים התירו את פרסום המידע.
3. אחריות כלפי לקוחות או מממנים – חסיון של כל מידע על אודות הלקוח, התוצרים והממצאים, למעט מידע שהלקוח התיר את פרסומו; מסירת מידע על תקני הפרסום לגורמים מעורבים במחקרים (שותפים, עמיתים, מממנים, לקוחות); מודעות למגבלות המומחיות והמיומנויות של עורכי המחקר באשר ליכולתם לבצע סוגי מחקר מסוימים.
4. אחריות כלפי הציבור – פרסום פומבי של שיטות והליכים שנעשו במחקרים שתוצאותיהם פורסמו לציבור; תיקון כל טעות שהתגלתה ושעשויה להשפיע על פרשנות הממצאים; תיקון כל ייצוג שגוי או הטיה בנתונים או בניתוח, לרבות אלו שנעשו על ידי שותפים, עמיתים, מממנים או לקוחות ויידוע של כל מי שנחשף לממצאים על הטעות ועל תיקונה.

³⁹ American Association for Public Opinion Research (2021, April). AAPOR Code of Professional Ethics and Practices. Retrieved from <https://aapor.org/standards-and-ethics/>.

5. אחריות כלפי המקצוע – מודעות לחשיבות הפרסום של רעיונות וממצאים מחקריים; מחויבות להתנהל בהתאם לכללי האתיקה של AAPOR, תוך ידיעה שאין בהם כל הכשר מקצועי לאדם או לארגון.
- ב. עקרונות של התנהלות מקצועית:
1. פיתוח שיטות לשם קבלת ממצאים מהימנים ותקפים – המלצה על כלים ושיטות לניתוח שמתאימים למטרות המחקר; זהירות מבחירה של כלים ושיטות לניתוח שעלולים להוביל למסקנות מטעות; זהירות מפרשנות מוטעית של ממצאים; זהירות בהצהרות לגבי מידת הביטחון של תוקף הממצאים והכללתם בהתאם למגבלות כגון מסגרת הדגימה; התבססות על נתוני אמת ודיוק בתיאור של מקורות המידע המצוטטים במחקר; לא להמציא או לזייף נתונים.
 2. שקיפות – תיאור של השיטות והממצאים בצורה מדויקת ומפורטת, בהתאם לתקני הפרסום של כללי אתיקה אלה.
- ג. תקנים למסירת מידע:
1. מידע שיש למסור בזמן דיווח הממצאים – אסטרטגית איסוף המידע (כגון סקר, קבוצת מיקוד), המממן והמבצע של המחקר, כלי המדידה, האוכלוסייה שנחקרה, שיטות להפקה ולגיוס של המדגם, שיטות האיסוף (כגון פנים אל פנים, סקר ממוחשב), מועדי איסוף המידע, גודלי המדגמים (לרבות טעות דגימה במדגמים הסתברותיים), שקלול המידע, אופן עיבוד המידע ובדיקות איכות, הצהרה כללית לגבי מגבלות המחקר.
 2. מידע נוסף שיש למסור לאחר דיווח הממצאים – למשל: הליכים שקשורים לניהול פנל (אם רלוונטי), שיטות לקידוד מידע לא-מספרי, חומרי עזר שנעשה בהם שימוש ועוד.
 3. גישה למאגרי מידע – יש לאפשר לחוקרים אחרים גישה למאגרי המידע, לאחר שהוסרו מהם מזהים אישיים, על מנת לאפשר שחזור ובדיקה בלתי-תלויה לתוקף הממצאים.
 4. תלונה על התנהלות לא תקינה – אם נפתחה חקירה רשמית על הפרת כללי האתיקה, יש לספק לחוקרים עמיתים מידע נוסף שיאפשר להם לבצע הערכה מקצועית של המחקר.

נוסף לאלה, ראוי לציין עקרונות אתיקה, ערכים וכללי יסוד למחקר ולהתנהלות של חוקרים, אשר פורסמו ופורטו על ידי האוניברסיטאות בארץ⁴⁰, להלן:

1. אמת – חיפוש אחר האמת המדעית ושאיפה להרחבת הידע, ההבנה והיכולת האנושיים.
2. חופש – תכנון וביצוע פעולות מחקריות ואקדמיות על בסיס יסוד חופש המחקר המדעי, בכפוף למגבלות מקצועיות, אתיות, משפטיות, סביבתיות וכספיות.
3. אחריות – החוקרים אחראים על כל היבטי המחקר, לרבות על בריאותם, בטיחותם, כבודם ופרטיותם של כל מי שעשוי להיות מושפע מהמחקר. על החוקרים להיזהר בשימוש בבעלי חיים במחקר ולשמור על הסביבה.
4. מקצועיות – התנהלות מחקרית בהתאם לסטנדרטים המקצועיים המקובלים הגבוהים ביותר, תוך דאגה לעדכון הידע העצמי כחוקרים והקניית ערכי היסוד וכללי המחקר המדעי לכל מי שמבצע מחקר תחת אחריות החוקרים.

40 להרחבה ודוגמאות ראו:

- אוניברסיטת בר-אילן – קוד אתי, אוניברסיטת בר אילן. אוחר מתוך
<https://acad-sec.biu.ac.il/sites/acad-sec/files/shared/Takanonim/Nehalim/KodEtti.19.2.2019.pdf>
- אוניברסיטת חיפה – הקוד האתי, מזכירות אקדמית, אוניברסיטת חיפה. אוחר מתוך
https://www.haifa.ac.il/wp-content/uploads/2014/03/code_of_ethics
עקרונות אתיקה בניהול מחקר במדעי החברה, ועדת האתיקה לאישור מחקרים, האוניברסיטה הפתוחה. אוחר מתוך
https://academic.openu.ac.il/education_psychology/ethics/pages/ethicscommittee.aspx
- קוד אתי, מכון ויצמן למדע. אוחר מתוך
https://www.weizmann.ac.il/wsos/PoliciesAndRegulations/sites/feinberg.PoliciesAndRegulations/files/procedures/h_0101_code_of_etics.pdf
- קוד אתיקה אקדמית, אוניברסיטת בן-גוריון בנגב. אוחר מתוך
<https://in.bgu.ac.il/acadsec/DocLib/Pages/ethics/>
- קוד התנהלות ראוייה במחקר, לשכת המשנה לנשיא למחקר, הטכניון, מכון טכנולוגי לישראל. אוחר מתוך
<https://research.technion.ac.il/files/2014/09/>
- קווים מנחים לעבודתה של ועדת האתיקה למחקרים לא קליניים בבני אדם, אוניברסיטת אריאל בשומרון. אוחר מתוך
<https://www.ariel.ac.il/wp/wp-content/uploads/2019/>
- תקנון אתיקה ונוהל מחקרים בבני אדם: נוהל ועדה מוסדית – אוניברסיטת רייכמן, אוניברסיטת רייכמן. אוחר מתוך
<https://www.runi.ac.il/media/algdsk2h/ethical-codes.pdf>
- תקנון בנוגע להתנהגות ראוייה במחקר, הרשות למחקר ופיתוח, האוניברסיטה העברית בירושלים. אוחר מתוך
https://openscholar.huji.ac.il/sites/default/files/mop/files/proper_conduct_code_6.pdf
- תקנון התנהגות ראוייה במחקר, רשות המחקר, אוניברסיטת תל-אביב. אוחר מתוך
https://www.tau.ac.il/sites/default/files/media_server/General/yoets/10-025.pdf

5. יושרה – ביצוע מחקר, ניתוח מידע ופרסומו בהתאם לדרישות ולסטנדרטים מדעיים, באופן ענייני, שקוף, מלא, מדויק, כן והוגן.

6. שיתוף פעולה – פעולת מחקר שבה שותפים מספר חוקרים, תיעשה תוך נכונות לעזור ואמון הדדי ותוך הכרה בתרומה אישית, קבוצתית ומוסדית של החוקרים.

נוסף לכך, יש המפרטים גם כללים להתנהלות מחקרית ראויה, לרבות אלה: שמירה על החוק, ביצוע מחקר ללא פניות וללא ניגוד עניינים, הימנעות ממחקרים שעלולים לכלול סיכונים, השתתפות במחקר מתוך חופש בחירה, שמירה על פרטיות המשתתפים ועל סודיות הנתונים, כיבוד קניין רוחני של אחרים, שקילת פרסום כשיש חשש לגרימת נזק ציבורי או אישי.

אוכלוסיות פגיעות או מוחלשות (vulnerable or disadvantaged) – קבוצות אוכלוסייה שנמצאות בסיכון ומוגבלות במימוש זכויותיהן עקב מגבלות או מחסומים בריאותיים (פיזיים, נפשיים, קוגניטיביים), כלכליים, חברתיים או פוליטיים. דוגמאות: אנשים עם מוגבלות, חוסים במעונות, מכורים לסמים, מהגרים, הורים יחידנים, קשישים, קטינים.

אוכלוסיית המטרה (target population) – קבוצת מרכיבים שעליהם החוקר מעוניין להסיק מסקנות באמצעות הסטטיסטיקה שנאספת מהמדגם. אוכלוסיית המטרה היא לרוב סופית בגודלה, היא תחומה בזמן ובמקום וניתנת למדידה.

אוכלוסיית הסקר (survey population) – עלולה להיות "צרה" יותר לעומת אוכלוסיית המטרה, עקב מגבלות שונות (למשל, לא ניתן לאסוף מידע בחלק מהאזורים). האוכלוסייה, בפועל, שממנה נאסף המידע בסקר.

אזור סטטיסטי (statistical area) – יחידת שטח רציפה שנוצרת מחלוקה גאוגרפית-סטטיסטית של יישוב שבו יותר מ-10,000 תושבים. האזורים הסטטיסטיים הם יחידות גאוגרפיות קטנות והומוגניות ככל האפשר, ובאמצעותם ניתן לשקף את המאפיינים הייחודיים של אזורים בתוך היישוב. אזור סטטיסטי מונה בדרך כלל 3,000-5,000 תושבים. זוהי היחידה הגאו-סטטיסטית הקטנה ביותר שלגביה מפורסמים נתונים סטטיסטיים רשמיים (אתר הלמ"ס).

אי-השבה (non-response) – פער בין המדגם המתוכנן ובין המדגם שענה בפועל על הסקר. מקור הפער עשוי להיות סירוב לענות, חוסר נגישות של נדגמים או אי-יכולת לענות. טעות אי-השבה עלולה לגרום להטיית אי-השבה (מותאם מספורטה ובייט-מרום, 2015).

אמידה סטטיסטית (statistical estimation) – הערכה של תכונות של אוכלוסייה באמצעות מניפולציות מתמטיות על הערכים הנדגמים ממנה.

אשכול – קבוצה של יחידות הנדגמות ביחד. במקרים רבים אשכול הוא יחידה טבעית של קבוצת יחידות כמו משק בית, בית ספר ואזור סטטיסטי. לדוגמה, בסקר משקי בית דוגמים משקי בית (אשכולות) וחוקרים את הנפשות שבהם (אתר הלמ"ס).

בדיקות מקדימות (questionnaire testing) – נעשות טרם איסוף הנתונים בשדה ומהוות שלב ראשון באיתור של הטיות מדידה. הבדיקות מתאימות בעיקר לשלבים הראשונים של פיתוח השאלון, הן איכותניות באופיין ואינן דורשות שימוש במדגמים גדולים.

⁴¹ חלק מההגדרות נלקחו או הותאמו ממקורות אלה - אתר הלשכה המרכזית לסטטיסטיקה (<https://www.cbs.gov.il/he/pages/default.aspx#>); ספורטה, ק. ובייט-מרום, ר. (2015). שיטות מחקר במדעי החברה: מונחון. תל אביב: האוניברסיטה הפתוחה.

גודל מדגם – מספר יחידות הדגימה ו/או יחידות החקירה במדגם (אתר הלמ"ס).

דגימה (sampling) – בחירה של חלק מיחידות האוכלוסייה כדי לקבל נתונים סטטיסטיים על האוכלוסייה. דגימה יכולה להיות הסתברותית או לא-הסתברותית.

דגימה הסתברותית (probabilistic sampling) – שם כולל להליכים לבחירת מדגם, שבהם היחידות שבמדגם נבחרות בצורה אקראית (רנדומלית) ממסגרת דגימה, ולכל יחידה יש סיכוי חיובי וידוע (לא בהכרח שווה) להיכלל במדגם.

דגימה בהסתברות יחסית לגודל (PPS - probability proportional-to-size sampling) – דגימה שבה הסיכוי של כל יחידה להיבחר הוא ביחס ישר לגודלה. לדוגמה, אם גודלה של יחידה א' הוא פי שלושה מגודלה של יחידה ב', הסיכוי של יחידה א' להידגם הוא פי שלושה מהסיכוי של יחידה ב' (אתר הלמ"ס).

דגימה לא-הסתברותית (non-probabilistic sampling) – שם כולל להליכים לבחירת מדגם, שבהם היחידות שבמדגם אינן נבחרות על בסיס הסתברותי.

דגימה רב שלבית (multistage sampling) – דגימה בשלבים, כאשר בכל שלב מוגדרות ונדגמות יחידות מתוך היחידות שהוגדרו ונדגמו בשלב הקודם. למשל, דגימת ישובים שמתוכם דוגמים בתי ספר שמתוכם דוגמים כיתות שמתוכם דוגמים תלמידים.

דגימת רשת קשרים (network sampling) – סוג של דגימה לא-הסתברותית, שמטרתה להגיע אל ו/או לחקור אוכלוסייה נדירה או חבויה.

הגדרה אופרציונלית (operational definition) – הגדרה הכוללת את כל האופרציות (פעולות) שיש לבצע כדי להפוך את המשתנה הנחקר ממונח תיאורטי למונח תצפיתי. היא מפרטת את כל הפעולות שנקט החוקר במחקר ספציפי כדי למדוד או להפעיל את המשתנה התיאורטי. הגדרה אופרציונלית ניסויית (experimental) מפרטת את התפעול שהחוקר מבצע בניסוי (כיצד הוא מתפעל או יוצר את ערכי המשתנה הבלתי-תלוי); הגדרה אופרציונלית נמדדת (measured) מפרטת את סוג המדידה שנעשית (כמו מבחן, שאלון או תצפית) כדי לגלות את ערכו של משתנה מסוים אצל נבדק מסוים (מותאם מספורטה ובייט-מרומ, 2015).

הגדרה מושגית, או נומינלית (nominal definition) – הגדרה מילונית של המונח (המשתנה) התיאורטי, המתארת אותו באופן כללי ומופשט (מותאם מספורטה ובייט-מרומ, 2015).

החזרת זיהוי (re-identification) – תהליך שבו מצליבים מידע עלום (אנונימי) עם מידע גלוי נגיש במטרה לגלות את היחידה (אדם, משק בית או ארגון) שאליו משתייכים הנתונים.

הטיה (bias) – סטייה של שיטת אמידה למאפיין של אוכלוסייה מערכו האמתי. הסטייה יכולה, למשל, לנבוע מאי-כיסוי של מסגרת הדגימה, מדגימה שאינה מקרית, מהבדלים בין משיבים ולא-משיבים, מחוסר דיוק במדידה או מחישוב שגוי של משקלות.

הטיית אי-השבה (non response bias) – הטיה הנוצרת כשהלא-משיבים (לו השיבו) היו נותנים תשובות השונות מאלה של המשיבים (מותאם מספורטה ובייט-מרום, 2015).

הטיית בחירה (selection bias) – הטיית בחירה היא סטייה של שיטת אמידה למאפיין של אוכלוסייה מערכו האמיתי, הנוצרת עקב הטיית דגימה ו/או גורמים שקשורים באיסוף המידע, למשל: קושי להגיע לקבוצות מסוימות, אי-השבה (סירוב מוחלט, הפסקת מענה או השבה חלקית), עצירה לא מתוכננת של איסוף הנתונים.

הטיית דגימה (sampling bias) – סטייה של שיטת אמידה למאפיין של אוכלוסייה מערכו האמיתי, הנובעת מדגימה שאינה מייצגת כראוי את האוכלוסייה.

הטיית כיסוי (coverage bias) – מצב שנוצר כאשר יש כיסוי חסר (under-coverage), כשחלק מהיחידות באוכלוסייה אינן נכללות במסגרת המדגם ו/או אינן מיוצגות במדגם, או במצב של כיסוי יתר (over-coverage), כשהמסגרת כוללת יחידות זרות או פסולות שאינן חלק מאוכלוסיית הסקר.

הטיית מדידה (measurement bias) – הטיה המתייחסת לכל שגיאה שיטתית או לא מקרית שמתרחשת במהלך איסוף הנתונים. מדידה מוטית עלולה להיגרם מכלי המדידה (למשל: ניסוח השאלות בשאלון), משיטת איסוף הנתונים, ממאפייני המשיב וממאפייני המראיין.

הפסקת מענה (break-off) – הפסקת מענה לסקר (להתראיין, להשיב לשאלון או להימדד בכל דרך אחרת) לאחר שהנדגם, או מראיין חליפי, נתן את הסכמתו להשתתף והתחיל במענה לסקר. הפסקת מענה נכללת בחישוב של שיעור סרבנות, שהוא אחד ממדדי התפוקה בסקר.

השבה (response) – מענה לסקר, באופן מלא או חלקי, לא כולל מי שהפסיק מענה לסקר ולא חידש אותו.

זקיפה (imputation) – תהליך השלמת נתונים חסרים, חסרי תוקף או לא מלאים, באמצעות מקורות חיצוניים או בשיטות סטטיסטיות (אתר הלמ"ס).

טעות דגימה (sampling error) – סטיית התקן של התפלגות הדגימה, שמתייחסת לסדר גודל של טעות צפויה באומד של מאפיין של אוכלוסייה. טעות זו מחושבת רק כאשר ניתן ליישם עקרונות הסתברותיים במערכת הדגימה.

טעות כיסוי (coverage error) – טעות הנובעת מחוסר הלימה בין אוכלוסיית המטרה לבין מסגרת הדגימה שלה. לחלק מאוכלוסייה אין סיכוי להיכלל במסגרת הדגימה (ספורטה ובייט-מרום, 2015).

טעות מדידה (measurement error) – טעות שיטתית שקשורה באיסוף הנתונים יכולה לנבוע ממספר גורמים: ניסוח השאלות בשאלון, שיטת איסוף הנתונים, מאפייני המשיב, מאפייני המראיין והקשר בין המראיין למשיב.

יחידת דגימה (sampling unit) – יחידה במסגרת הדגימה, שעבורה נקבע אם היא נדגמת או לא.

יחידת חקירה (statistical unit) – יחידה שנחקרת. לעתים יחידת החקירה שונה מיחידת הדגימה, למשל במצב שדירת מגורים היא יחידת הדגימה והנפשות בדירה הן יחידות החקירה.

יצירת קשר (contact) – השגת הנדגם (למשל: באמצעות פנייה פנים אל פנים, בדואר רגיל או אינטרנטי וכיו"ב). כישלון ביצירת קשר (non-contact) מורכב ממצבים שבהם עורכי הסקר לא הצליחו ליצור קשר.

מדגם (sample) – תת-קבוצה של האוכלוסייה, שנבחרה לשם קבלת מידע על האוכלוסייה שבה מתעניינים (אתר הלמ"ס).

מדגם אשכולות (cluster sample) – מדגם הסתברותי שבו אשכולות הטרוגניים, לרוב טבעיים (כגון משק בית, בית ספר ואזור גיאוגרפי), נבחרים באופן מקרי מהאוכלוסייה. לאחר דגימת האשכולות בודקים את כל היחידות שבכל אשכול או שממשיכים לדגום מהאשכולות.

מדגם ברוטו (gross sample) – המדגם כפי שהתקבל, כולל משיבים, לא-משיבים (נפל) ולא-שייכים לאוכלוסיית הסקר (מקרי אפס) (אתר הלמ"ס).

מדגם הסתברותי (probabilistic sample) – התוצאה של דגימה הסתברותית.

מדגם לא-הסתברותי (non-probabilistic sample) – התוצאה של דגימה לא-הסתברותית.

מדגם מותאם (sample matching) – מדגם לא-הסתברותי בו הרכב היחידות מותאם לאוכלוסייה באמצעות השוואה לנתוני מפקד, נתונים מנהליים, או מדגם הסתברותי, בהתאם למשתני החקירה ומשתני עזר הרלוונטיים לסקר.

מדגם מייצג (representative sample) – מדגם המשקף במשתנים הרלוונטיים את האוכלוסייה שאליה החוקר רוצה להכליל את מסקנות המחקר (ספורטה ובייט-מרום, 2015).

מדגם מכסה (quota sampling) – מדגם לא-הסתברותי הנחשב כמדגם נוחות, או שהוא סוג של דגימה מותאמת, אם הוא כולל מכסות כלכליות-חברתיות אשר משקפות התפלגויות של משתנים באוכלוסייה שהם משמעותיים לסקר. המדגם "מעוצב", כך שהרכבו ייצג את הרכב האוכלוסייה במשתנים הרלוונטיים.

מדגם מקרי (או אקראי) פשוט (SRS – simple random sample) – מדגם הסתברותי בו היחידות נבחרות בצורה שמבטיחה לכל יחידה אותה הסתברות להיכלל במדגם (אתר הלמ"ס).

מדגם נוחות (convenience sample) – מדגם לא-הסתברותי, שהשיקול העיקרי בו הוא נוחות האיתור או גיוס המשתתפים לסקר.

מדגם נטו (net sample) – מדגם המשיבים, ללא מקרי נפל ומקרי אפס (אתר הלמ"ס).

מדגם שכבות (stratified sample) – מדגם הסתברותי בו היחידות נבחרות באופן מקרי מכל שכבה שהוגדרה כתת-אוכלוסייה הומוגנית, כך שכל יחידה באוכלוסייה תשתייך לשכבה אחת בלבד.

הדגימה יכולה להיות פרופורציונלית לגודל השכבה באוכלוסייה או לא פרופורציונלית, למשל במקרה של קבוצות קטנות מאוד, אשר לאחר מעשה מקבלות תיקון למשקלן האמתי באוכלוסייה.

מדגם שיטתי (systematic sample) – מדגם הסתברותי בו היחידה הראשונה נדגמת באופן מקרי ושאר היחידות נדגמות בדילוגים קבועים, לפי יחס נתון. ניתן להתייחס לדגימה זו כאל דגימה מקרית, אם מסגרת הדגימה אינה מסודרת ביחס למשתנים הנחקרים.

מדד (index) – תוצר שקלול של פריטים (כגון שאלות בשאלון) שמודדים משתנה תיאורטי מורכב (מותאם מספורטה ובייט-מרום, 2015).

מדד חברתי-כלכלי (socio-economic index) – מדד המאפיין יחידות גאוגרפיות לפי הרמה החברתית-כלכלית של אוכלוסיית התושבים. הרמה החברתית-כלכלית של האוכלוסייה נמדדת באמצעות שילוב של תכונותיה הבסיסיות בתחומים הבאים: הרכב דמוגרפי, השכלה וחינוך, רמת חיים, תעסוקה וגמלאות. ערך המדד מחושב כערך רציף ומתוקן. ערך חיובי של המדד מצביע על כך שהרמה החברתית-כלכלית של האוכלוסייה ביחידה הגיאוגרפית גבוהה מהממוצע. אשכול חברתי-כלכלי הוא תוצאה של חלוקת היחידות הגאוגרפיות לקבוצות הומוגניות לפי ערך המדד שלהן (מותאם מאתר הלמ"ס).

מהימנות של כלי המדידה (reliability) – אמת מידה לאיכות המדידה, המשקפת את מידת העקביות או היציבות של המדידה. מקדם המהימנות מבטא את היחס בין השונות האמיתית ובין השונות הנצפית של המדידה (שכוללת שונות אמיתית ושונות טעות). מהימנות המדידה נבדקת לרוב באמצעות מהימנות פנימית (internal reliability) ומהימנות מבחן חוזר (test re-test reliability).

מהימנות מבחן חוזר (repeatability or test-retest reliability) – מידת היציבות של המדידה לאורך זמן, אשר משקפת את מידת הרגישות של כלי המדידה לטעויות מקריות שנובעות מגורם הזמן. מקדם המהימנות מחושב על ידי מתאם בין שתי מדידות שנעשות באותו כלי המדידה, בשני מועדים שונים לאותם נבדקים. ככל שהמתאם גבוה יותר כך כלי המדידה יציב יותר ורגיש פחות לשינויים בגורמים הקשורים בזמן המדידה (מותאם מספורטה ובייט-מרום, 2015).

מהימנות פנימית (internal reliability / consistency) – מידת העקביות הפנימית בין שאלות, שנועדו למדוד את אותו תחום מהיבטים שונים. מהימנות פנימית משקפת את מידת הרגישות של כלי המדידה לטעויות מקריות שנובעות מתוכן השאלות. מקדם המהימנות מחושב על ידי ממוצע המתאמים בין כל פריט לכל הפריטים האחרים, כשהנפוץ הוא מקדם אלפא של קרונברך (Cronbach's alpha). ככל שמקדם המהימנות גבוה יותר כך הקשר בין השאלות באותו התחום נחשב חזק יותר (מותאם מספורטה ובייט-מרום, 2015).

מזהה ישיר (direct identifier) – מאפיין ייחודי של נדגם, שמאפשר לזהות את הנדגם באופן חד-חד-ערכי. דוגמה: מספר תעודת זהות.

מזהה עקיף (indirect identifier) – מאפיין של נדגם, שהוא לעצמו אינו מאפשר לזהות את הנדגם באופן בלבדי, אך צירוף של מספר מזהים עקיפים, או שילובם עם מידע ממקור אחר עלול לאפשר החזרת זיהוי (re-identification). דוגמאות: גיל, השכלה, כתובת, מצב בריאות.

מידע אישי שניתן לזיהוי (personally identifiable information) – מידע שמאפשר לזהות נדגמים, באמצעות מזהה ישיר (direct identifier), כגון שם ומספר תעודת זהות, או באמצעות מזהה עקיף (מכונה גם מזהה למחצה; indirect identifier), כגון גיל והשכלה, שהם לעצמם אינם מזהים אך צירופם עלול לאפשר החזרת זיהוי (re-identification).

מידע חסר (missing data) – חוסר במידע עשוי להיות ברמת היחידה (למשל: שאלון, תיק רפואי ועוד) או ברמה של פריט מידע מתוך היחידה השלמה שנבדקת (למשל: תשובה לשאלה בשאלון, תוצאה של בדיקה רפואית ועוד). מידע חסר ברמת היחידה יכול לנבוע מאי-השבה כיוון שלא נוצר קשר עם נדגם או שהוא סירב לשתף פעולה, או שהמידע אבד או נפגם במהלך האיסוף או העיבוד. מידע חסר ברמת הפריט יכול לנבוע משיתוף פעולה חלקי, מסיבות טכניות (למשל: אי-רציפות בקליטת נתונים במוניטור) או מסיבות אחרות לא ידועות (למשל: חוסר במידע בתיק רפואי של נבדק).

מידע נלווה (paradata) – מידע שמתקבל במהלך התקדמות הסקר ומתאר או מגדיר את רכיבי המידע ואת השיטות והתהליכים שקשורים באיסוף המידע ובעיבודו. למשל: פרטים שנוגעים לשיטת איסוף המידע (mode); לטיפול בחוסר מענה, סירובים והפסקות מענה; למידע על הפרעות באיסוף המידע והתנהגות חריגה של מראיינים; למידע על משיבים ועל לא-משיבים ועוד.

מידע עילום (anonymous data) – מידע שאין אפשרות לזהות נדגם כלשהו באמצעות הנתונים, או שהסיכוי לכך הוא נמוך ביותר. המידע יכול להיות עלום מלכתחילה, כיוון שלא נאספו כל נתונים מזהים, או לאחר מעשה, כיוון שהוא עבר תהליך של עילום (anonymization, התממה).

מסגרת דגימה (sampling frame) – רשימת כל היחידות המתאימות להגדרת אוכלוסיית הסקר, שממנה מוציאים מדגם. לדוגמה, בסקר כוח אדם, רשימות הארנונה משמשות כמסגרת לדגימת הדירות ביישובים העירוניים (אתר הלמ"ס).

מערך מחקר (research design) – המבנה, התכנון והאסטרטגיה של דרך איסוף הנתונים וניתוחם. במערך מחקר מתאמי (correlational) נערכת תצפית במשתנים, בעוד שבמערך ניסויי (experimental) ודמוי ניסויי (quasi-experimental) נעשה תפעול של המשתנה הבלתי-תלוי בקבוצת הניסוי (אך לא בקבוצת הביקורת) על מנת לבדוק את השינוי שחל במשתנה התלוי (מותאם מספורטה ובייט-מרום, 2015).

מפקד (census) – תהליך של איסוף, קליטה, הערכה, ניתוח, פרסום והפצה של נתונים המתייחסים לאוכלוסייה שלמה בנקודת זמן נתונה.

מקרה אפס (non-eligible case) – יחידה שנדגמה לסקר אך אינה שייכת לאוכלוסיית הסקר (אתר הלמ"ס).

מראיין חליפי (proxy respondent) – משיב בשם הנדגם, למשל: הורה שמשיב במקום קטין, מטפל שמשיב במקום מטופל שאינו יכול להשתתף בסקר. במקרים מסוימים בן משק בית משמש כמראיין חליפי כשהוא מוסר פרטים על בני משק הבית האחרים.

משק בית (household) – אדם אחד או קבוצת אנשים הגרים יחד בדירה אחת באופן קבוע ברוב ימות השבוע, ויש להם תקציב הוצאות משותף למזון. הם עשויים להיות קרובי משפחה, לא קרובי משפחה או צירוף של אנשים קרובים ושאינם קרובים (אתר הלמ"ס).

משקל (weight) – מספר היחידות באוכלוסיית הסקר המיוצגות על ידי היחידה הנדגמת או היחידה הנחקרת (אתר הלמ"ס).

משתנה עזר (auxiliary variable) – גורם משותף שקשור הן למשתנה הנחקר והן לנטייה להשיב (או לא להשיב). המשתנה מסייע להתמודד עם מקורות אפשריים להטיה – הטיית דגימה והטיית אי-השבה – באמצעות יצירת מדגם מותאם, שקלול של נתוני המדגם וזקיפה של ערכים.

מתודולוגיה מכוונת תוצאה (preempted outcomes) – מתודולוגיה שיש בה התניה מראש לתוצאות, באמצעות בחירה ו/או התאמה של השיטות לשם קבלת הממצאים המצופים. בחירת השיטות או התאמתן עשויה להתבצע בכל שלב בסקר, לרבות בהגדרת המדגם, בקביעת השאלות בשאלון, באופן הפנייה לנדגמים ואיסוף הנתונים, בדרכי העיבוד והניתוח, באופן הצגת הממצאים ובפרשנות לממצאים.

נדגם (sampled unit) – אובייקט או יחידה שנבחרו מתוך האוכלוסייה ואשר מהווים חלק מהמדגם.

נחקר (respondent) – נדגם שהמאפיינים ו/או התגובות שלו נמדדים בסקר.

נטל השבה (response burden) – המידה שבה משיב לסקר תופס את השתתפותו בסקר כקשה, גוזלת זמן או מעיקה מבחינה נפשית. נטל השבה מושפע מגורמים אובייקטיביים כגון אורך הראיון ומורכבות המטלה הנדרשת מהמשיב ומגורמים סובייקטיביים כגון תחושתו של המשיב ביחס לקושי, לזמן ולמאמץ שנדרש ממנו ותחושת לחץ כתוצאה משאלות שנתפסות על ידו כפולשניות.

ניתוח סדרתי (sequential analysis) – דגימה הסתברותית בה נתונים מתקבלים באופן סדרתי והדגימה מופסקת לפי פרוטוקול מוגדר מראש, כזה שלא מגדיר מראש את גודל המדגם.

סולם מדידה (measurement scale) – מערכת מספרית שנבחרת לייצג אובייקטים או אירועים. סולם מדידה יכול להיות אחד מארבעה: שמי (שערכיו מבטאים את זהות האובייקט, כגון "ישוב עירוני", "ישוב כפרי", "ישוב אחר"); סדר (סולם שמי עם מדרג בין הערכים, כגון "השכלה יסודית", "השכלה תיכונית", "השכלה אקדמית"); רווח (סולם סדר עם הפרשים ידועים או שווים בין הערכים,

כגון מד טמפרטורה); מנה (סולם רווח עם נקודת אפס קבועה ומוחלטת, כגון גובה או משקל) (מותאם מספורטה ובייט-מרום, 2015).

סיכון (risk) – מדד למידת האיום על ידי נסיבות או אירוע. בסקר עלולים להיות סוגים שונים של סיכון לנדגמים ולמשתתפים, כגון פיזיים (למשל: פגיעה גופנית), נפשי (למשל: חווית רגש שלילי), חברתיים (למשל: איבוד מעמד חברתי) וחוקיים (למשל: דליפת מידע אישי, חוסר מודעות לזכויות, גילוי עצמי על התנהגות לא חוקית). סיכונים אחרים יכולים להיגרם לפעולה הסטטיסטית, כגון מתודולוגיים (למשל: בחירה של דגימה לא מתאימה), טכניים (למשל: כשל במערכות ממוחשבות לאיסוף הנתונים) או אתיים (למשל: בחירת מתודולוגיה מכוונת תוצאה).

סקר (survey) – שיטה לאיסוף נתונים על תופעה מסוימת כפי שהיא מתרחשת בסביבתה הטבעית, באמצעות מדגם. הסקר מניב מסד נתונים מובנה ומסודר המתקבל ממחקר מתאמי וממנו מוכללים הממצאים לכלל האוכלוסייה באמצעים סטטיסטיים (מותאם מספורטה ובייט-מרום, 2015).

סקר מקדים (מבחן מקדים, pretest) – בדיקה אמפירית, כמותית, של טיוטת השאלון במדגם קטן מאוכלוסיית המטרה. מטרת הבדיקה היא לזהות קשיים בהשבה על השאלון ובהליך איסוף הנתונים ולהתמודד איתם, בטרם עריכת הסקר (מותאם מספורטה ובייט-מרום, 2015).

סרבנות (refusal) – אי-הסכמה להיחקר (להתראיין, להשיב לשאלון או להימדד בכל דרך אחרת) לאחר שנוצר קשר עם הנדגם או עם מראיין חליפי.

פנל (panel) – מאגר משיבים פוטנציאליים, שהסכימו להשתתף או שפנו להשתתף בסקרים בנקודות זמן שונות, לעתים עבור תשלום או עבור הטבות עם שווי כספי. הפנל יכול להיות הסתברותי או לא-הסתברותי ודרך איסוף הנתונים עשויה להיות באמצעות האינטרנט, הטלפון, הדואר, פנים אל פנים, או כל שילוב של דרכים, בהתאם למטרות הסקר.

פנל הסתברותי (probability panel) – משתתפים שנדגמו מתוך מסגרת דגימה ושנתנו את הסכמתם להשיב על סקרים, במשך תקופת זמן, לעתים עבור תשלום או עבור הטבות עם שווי כספי.

פנל לא-הסתברותי (non-probability, online or opt-in panel) – משתתפים שפנו להצטרף לפנל ולהשיב על סקרים, או שנתנו את הסכמתם לכך, לעתים עבור תשלום או עבור הטבות עם שווי כספי.

קבוצת אוכלוסייה (population group) – חלק מהאוכלוסייה, מוגדר לפי מאפיין רקע.

רצייה חברתית (social desirability) – צורך של משתתף במחקר להיראות בסקר לפי אמות המידה המקובלות בקבוצת ההתייחסות שלו, דבר שעלול להוביל אותו להגיב באופן מוטה, לדוגמה: לא להודות בהתנהגות או בעמדה שאינה מקובלת (מותאם מספורטה ובייט-מרום, 2015).

שאלות במבנה לוח (matrix questions) – שאלות שסולם התשובות בכולן זהה ולכן מארגנים אותן בלוח (טבלה), מתוך הנחה שכך נוח יותר להשיב עליהן (מותאם מספורטה ובייט-מרום, 2015).

שאלת סינון / ניתוב (filter question) – שאלה שמבחינה בין תת-שאלונים בתוך שאלון נתון, למשל: שאלות שמיועדות רק לאנשים עובדים, שאלות שמיועדות רק לאנשים עם מוגבלות וכדומה.

שבר דגימה (sampling fraction) – היחס בין גודל המדגם המתוכנן לבין גודל האוכלוסייה (אתר הלמ"ס).

שיטת אמידה (estimation method) – שיטה להערכת מאפיין של אוכלוסייה.

שיטת דגימה (sampling method) – שיטה להוצאת מדגם (אתר הלמ"ס).

שיטת זקיפה (imputation method) – שיטה להשלמת נתונים חסרים, חסרי תוקף או לא מלאים (מותאם מאתר הלמ"ס).

שיעור גיוס (recruitment rate) – מדד תפוקה בפנל הסתברותי, המבטא את היחס בין מספר הנדגמים שנתנו את הסכמתם להשתתף בפנל ובין מספר המקרים הכללי במדגם.

שיעור השבה (response rate) – מדד תפוקה מרכזי בסקר, המבטא את היחס בין כל המקרים שהשיבו באופן מלא או חלקי לסקר ובין כלל המקרים במדגם.

שיעור השבה כולל (cumulative response rate) – מדד מורכב שמשמש בפנל הסתברותי, שבא להעריך את איכות הפנל; מכפלה של שיעור הגיוס, השלמת שאלון פרופיל והשלמת שאלון הסקר. אם מתקיימים מספרי גלי בדיקה המכפלה תכלול גם את שיעור ההתמדה (כלומר: מספר המשיבים בגל עוקב מתוך מספר המשיבים בגל הקודם).

שיעור השלמת פרופיל (profile rate) – מדד תפוקה בפנל, המבטא את היחס בין מספר הנדגמים שמילאו את פרטי הרקע שלהם בשאלון הפרופיל (המשמשים לסינון נדגמים בהתאם לנושא הסקר ו/או לעיצוב של מדגם מותאם) ובין כלל הנדגמים שנתנו את הסכמתם להשתתף. מחושב עבור פנלים הסתברותיים ולא-הסתברותיים.

שיעור השלמת שאלון (completion rate) – מדד תפוקה בפנל, המבטא את היחס בין מספר המקרים שבהם התקבל מענה לשאלון של סקר מסוים ובין מספר חברי הפנל שמלאו פרטי פרופיל. (הערה: ההנחה היא שרק מי שמילא פרטי פרופיל הוא מועמד למענה לשאלון).

שיעור יצירת קשר (contact rate) – מדד תפוקה בסקר, המבטא את היחס בין כל המקרים שבהם נוצר קשר עם בן משק בית אחראי ובין כלל המקרים במדגם.

שיעור סרבנות (refusal rate) – מדד תפוקה בסקר, המבטא את היחס בין כל המקרים שבהם היה סירוב לענות לסקר או שהייתה הפסקת מענה לסקר ובין כלל המקרים השייכים לאוכלוסיית הסקר שאי פעם נוצר איתם קשר.

שיעור שיתוף פעולה (cooperation rate) – מדד תפוקה בסקר, המבטא את היחס בין כל המקרים שנחקרו (באופן מלא או חלקי) ובין כלל המקרים השייכים לאוכלוסיית הסקר שאי פעם נוצר איתם קשר.

שיתוף פעולה (cooperation) – השתתפות בסקר באופן מלא או חלקי, בהתאם לאמות מידה שנקבעות בסקר, למשל: מענה על אחוז מסוים משאלות הסקר שהן רלוונטיות לנדגם ו/או מענה על אחוז מסוים משאלות הליבה בסקר.

תוקף המדידה (validity) – אמת מידה לאיכות המדידה; מבטא את המידה שבה כלי המדידה מודד את מה שהוא אמור למדוד, כלומר את ההלימה בין המשתנה התצפיתי (שנמדד) למשתנה התיאורטי. נהוג למדוד תוקף תוכן (content validity), תוקף מבנה (construct validity) ותוקף תלוי קריטריון (criterion related validity).

תוקף חיצוני (external validity) – משקף את המידה שבה ניתן להכליל את ממצאי הסקר על אוכלוסיות, מצבים וזמנים שונים.

תוקף מבנה (construct validity) – משקף את המידה שבה כלי המדידה מצליח לתרגם תפיסה תיאורטית של מושג, רעיון או התנהגות לרכיבים אופרציונליים, מציאותיים. תוקף מבנה מבטא את מידת ההלימה של התפיסה התיאורטית עם המדד האופרציונלי.

תוקף פנימי (internal validity) – עוסק בשאלה אם הטיפול שניתן למשתנה הבלתי-תלוי גרם לשינוי במשתנה התלוי (בקבוצת הניסוי) ולא משתנה חיצוני כלשהו (ספורטה ובייט-מרום, 2015).

תוקף תוכן (content validity) – משקף את מידת הרלוונטיות של התחומים שנועדו להימדד ואת הייצוג שלהם באמצעות פריטי שאלון. תוקף תוכן נבדק בשלבים ראשונים של פיתוח השאלון.

תוקף תלוי קריטריון (criterion related validity) – מבטא את הקשר בין מידע שהופק מכלי המדידה ובין נתונים עובדתיים שניתן לסמוך שהם מהימנים ותקפים.

תמריץ (incentive) – לרוב תגמול כספי או בעל ערך כספי (כגון שובר קנייה או מתנה), אך הוא עשוי להיות בעל ערך למשיב, אם כי ללא ערך כספי (למשל: דוח סיכום אישי לארגון על ממצאי הסקר, בהשוואה לארגונים אחרים). התמריץ ניתן במטרה להעלות את שיעורי ההשבה לסקר והוא יכול להינתן מראש, או להיות מותנה בהשתתפות בסקר.