

האם סקר העמדות הארגוני שלך מהימן?

אבי בניטה, יועץ ארגוני וסטטיסטי

www.avibenita.com

רקע

כל תלמיד שיטות מחקר לומד שאיכותו של סקר, או כל כלי מדידה אחר, נקבעת עפ"י מידת **התקפות** (היכולת לשקף את הנושא הנמדד) ו**המהימנות** שלו (היכולת להיות עקבי כך ששני עובדים עם עמדה זהה כלפי שאלה מסוימת, יעניקו לה ציון זהה כל זמן)

בעוד שתקפות קשה להערכה, מאחר והיא מחייבת קיומם של מדדים "אובייקטיביים" נוספים מולם ייבחנו תוצאות הסקר, נוח יותר יחסית לטפל בסוגיית **המהימנות** מאחר והיא ברובה תלויה באופן בניית השאלון וניתוח הנתונים

סקר לא מהימן עשוי להיות לא רק חסר ערך אלא גם בעל השלכות שליליות מאחר והוא יוצר תמונת עולם מעוותת של הנושא הנמדד ועלול להוביל לנקיטת פעולות שגויות ומזיקות. ייתכן שמבין כל הגורמים הקובעים איכותו של הסקר, נושא המהימנות הנו הקריטי ביותר

למרות הניחוח האקדמי שלה, סוגיית המהימנות הנה לגמרי מעשית ונוגעת למהות של הסקר ולא להיבט אזוטרי שלו

מפליא לכן לראות שהנושא אינו זוכה בלשון המעטה לתשומת הראויה ע"י העוסקים בתחום. באופן כללי, ניתן לומר שהעיסוק והמודעות לסוגיות מתודולוגיות בתחום סקרי העמדות הארגוניים, הנו דל מאוד וביחס הפוך לפופולריות של תחום הסקרים בשנים האחרונות

יש סיבות רבות לכך, אבל המשמעותית ביותר להערכתנו טמונה בקושי להבדיל במבט ראשון בהבדל בין סקר מהימן/לא מהימן. כניסתן של המערכות האינטרנטיות "האוטומטיות" החריפו עוד יותר את המצב מאחר והדגש עבר מהתוכן והניתוח לעיבוד והצגה, ובדרך הלכה קצת לאיבוד "הנשמה" של הסקר

מטרה

מאמר זה סוקר מקצת מהסוגיות העיקריות המשפיעות על מהימנות הסקר והדרכים האפשרויות לטפל בהן. הוא נכתב מנקודת מבט מאוד יישומית ופחות אקדמית. הוא מתבסס על ניסיון בין מאות סקרים בארגונים ומבטא את דעתו של כותב המאמר. חלק מהסוגיות במאמר זה נסקרו יותר בקצרה [במאמר קודם](#) אודות פרדיגמות בסקרי עמדות

למרות שנושא המהימנות ישים לכל סקר או כלי מדידה, מאמר זה מתייחס לארבעה סוגי סקרים אשר כל ארגון מצוי מבצע מעת לעת במינן כזה או אחר:-



המאמר בוחן את סוגיית המהימנות עפ"י שני האשכולות העיקריים בהם היא באה לידי ביטוי:-





סוגיות הקשורות לשאלון עצמו

מיותר לציין עד כמה איכותו של השאלון קריטית. שאלון הבנוי היטב יספק תוצאות בעלות ערך גם אם ניתוח התוצאות הנו בסיסי ופשוט. לעומת זאת, הסטטיסטיקה המתקדמת ביותר לא תושיע נתונים המתבססים על שאלון עם בעיות מהימנות חמורות

אופן ניסוח השאלות

מדובר בסוגיה המהווה עקב אכילס של שאלונים רבים. היא נובעת מכך שמחבר השאלון ממוקד יותר במוצר (מה לשאול?) ופחות בלקוח, קרי ממלא השאלון (איך לשאול?)

כך מתקבלות פעמים רבות שאלות בלתי בהירות או לא מספיק חדות, הניתנות לפרשנויות רבות. הציון הניתן עקב כך לשאלה מסוימת תלוי יתר על המידה באופן בו היא הובנה, דבר הפוגע כמובן במהימנות של התוצאה מאחר ושני עובדים בעלי עמדה זהה כלפי הנושא שנשאל עליו, עלולים לתת ציון שונה לאותה שאלה בגלל שהיא התפרשה אחרת על ידם

להלן דוגמאות מוחשיות (אמיתיות) אשר עשויות להמחיש את הטעויות הנפוצות ביותר:-

השאלה	מה עלול לפגוע במהימנות?
"אני זוכה להכרה ולהערכה על עבודתי/הישגי"	"עבודתי" ו"הישגי" הם שני דברים שונים. אז למה הכוונה? לא ברור מי אמור לספק את ההערכה וההכרה (המנהל הישיר? הצוות? משאבי אנוש?...) "הכרה ולהערכה": למה גם וגם? ואולי בכלל אין לממלא השאלון הישגים יוצאי דופן, ולכן הוא לא זכאי להערכה (ולא בגלל שבחברה שלו לא מעריכים הישגים)
"יחידת הרכש מקפידה על עמידה בהתחייבויות, בלוחות זמנים ופועלת לממש את המטלות שלה"	מיותר לציין עד כמה השאלה לא ממוקדת וממחישה את האימרה לפיה "תפשת מרובה, לא תפשת..."
"In my team, there is a clear sense of identity and belonging"	מעבר לסוגיית כפל הביטויים (identity and belonging), מטרידה יותר המידה בה העובד הממוצע מבין למה הכוונה? הנושא בהחלט ראוי אך הוא עשוי להיות מובן, אם בכלל, בצורה שונה ע"י עובדים שונים
"אני שותף בקבלת החלטות בדברים הקשורים לעבודתי"	השאלה מניחה שצריך לשתף כל עובד בכל החלטה הקשורה לעבודתו. האמנם? אולי בכלל העובד המצוי אינו מצפה ואף אינו שואף להיות שותף להחלטות, ומניח שזו עבודתם של מנהלים?
"אני מרגיש שיש לי מוטיבציה ומחויבות כלפי מקום עבודתי"	האם מישהו מצפה שהעובד יצהיר שאין לו מוטיבציה ומחויבות??? לא מפליא לכן ששאלות המנוסחות בצורה זו זוכות תמיד לציונים גבוהים מאוד ולשונות מאוד נמוכה נושא מסוג זה עדיף לשאול בצורה השלכתית בסגנון של "העובדים בצוות שלי בעלי מוטיבציה ומחויבות כלפי מקום עבודתם" הניסיון מראה שהציונים נבדלים בצורה חדה בין 2 הניסוחים

רוחב סולם התשובות

דילמה הקשורה לסולם התשובות, נוגעת למספר הדרגות שלו. עוסקים לא מעטים בתחום הסקרים נוטים להשתמש בסולם בין 10 דרגות מתוך אמונה שהוא יותר רגיש מסולם היותר מקובל בין 5 דרגות, ולכן הוא צפוי לייצר שונות גבוהה יותר בתשובות. נימוק נוסף של חסידי הסולם המורחב הוא

2. To what extent do these attributes fit your experience with the Service Provider on a scale from 1 (Very negative) to 10 (Very positive)?

	1	2	4	5	6	7	8	9	10
Demonstrates expertise and proficiency	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>

האינטואיטיביות שלו מאחר והוא מזכיר את סולם ציוני בית הספר ותוכניות הריאליטי

החיסרון הגדול בסולם בין 10 דרגות טמון בכך שלא ניתן לייחס מלל מפורש לכל דרגה בסולם, ולכן אין כל אפשרות להיות בטוח שמי שענה 7 אכן פחות מרוצה ממי שענה 8, דבר העשוי כמובן לפגוע באופן חמור במהימנות של התוצאות

בסולם בין 5 דרגות לעומת זאת, ניתן לקשור ביטוי מילולי לכל דרגה בסולם, ולמעשה ממלא השאלון

3. To what extent do these attributes fit your experience with the Service Provider?

	To a very high extent	To a high extent	To a moderate extent	To a low extent	To a very low extent
Demonstrates expertise and proficiency	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

אינו נתן ציון אלא בוחר אמירה מבין חמשת האמירות המוצעות לו

חיסרון נוסף של סולם רחב מידי טמון ביכולת המוגבלת לייצג התפלגות של הציונים, ולמעשה רוב המשתמשים בו עוקפים את המגבלה ע"י כך שהם מצמצמים אותו באופן מלאכותי ל-5 דרגות. אם כך, למה מלכתחילה היה צריך סולם רחב?

הטיעון של רגישות הסולם בגלל ריבוי הדרגות בו הנו די גבולי מאחר ולמעשה לא מתקבלות תשובות בדרגות הנמוכות מאוד. עשויה בכל זאת להתקבל שונות גבוהה יותר מאשר בסולם בין 5 דרגות אך מדובר לדעתנו בניפוח מלאכותי הנובע מכך שהבחירה בין דרגות צמודות (כגון 8 או 7) הנה די אקראית

יתירה מזו, מרבית פרסומי ה-Benchmark בעולם הם על סולם בין 5 דרגות. סיבות אלה ואחרות (כגון הקושי בהצגת התפלגות בת 10 קטגוריות), מביאות אותנו להמליץ על סולם תשובות בין 5 דרגות על פני סולם בין 10 דרגות

לסולם בן 10 דרגות עשויים להיות יתרונות בסקרים טלפוניים ארוכים בהם קשה להציג את הסולם המילולי למשיב ולכן נותנים לא לבחור ציון בין 1 ל-10 מתוך הנחה שאין לכך השלכות מכריעות מאחר ומדובר בד"כ בשאלונים בעלי אופי שיווקי ולא סקרי עמדות ארגוניים בעלי השפעה על ניהול ההון האנושי בחברה

סוגיות הקשורות לניתוח הסטטיסטי

סוגיית מיצוע השאלות

פגיעה במהימנות עלולה לנבוע גם מאופן ניתוח התוצאות. העיקרית שבהן קשורה לאופן בו מקבצים שאלות שונות לנושא אחד

שאלונים רבים בנויים מנושאים שונים כשכל נושא מכיל מספר שאלות. בשאלון אקלים ארגוני, נושא כזה עשוי להיות לדוגמה "סביבת העבודה" אשר יכולה שאלות בתחום המחשוב, האחזקה, נוחות מקום הישיבה ואולי גם רמת האוכל והיחס עם העמיתים בצוות

נראה לכן די טבעי שבניתוח הסקר, ממצעים כל השאלות ויוצרים ציון ממוצע וכולל של "סביבת העבודה". יש כאלה שמרחיקים לכת עוד יותר ומחשבים ציון ממוצע לכל השאלות בסקר ומולו משווים את הציונים של השאלות/תחומים שונים. ציון כזה משמש אח"כ כסוג של אינדקס מוביל אתו מתבצע בנצ'מרק והשוואה בין יחידות וחישוב מגמות לאורך השנים

נוהג כזה **פסול לחלוטין מבחינה סטטיסטית** מאחר והוא למעשה "מערבב תפוזים ועגבניות" ויוצר תוצאה מאוד לא מהימנה. מה מייצג לדוגמה ממוצע הציון של שביעות רצון מרמת ניקיון בחדרים לבין שביעות רצון מהתקשורת הארגונית?

לא ניתן כאן להסביר לעומק את הבעיה הטמונה בגישה הנ"ל, אך היא קשורה באופן כללי לכך שאסור למצע שאלות בעלות מתאם נמוך ביניהן, כלומר שאינן מייצגות את אותו עולם תוכן

ככלל, נדרש לעשות בדיקות מקדימות לפני שמחליטים לאגד מספר שאלות לממד אחד. המומלצת ביניהן היא טכניקת ניתוח הגורמים (Factor Analysis) אשר מסייעת לאתר נושאים וממדים בעלי מכנה משותף אשר ניתן להכליל בתוך ממד משותף. חשוב לזכור שגם ניתוח גורמים הנו כלי תומך ובשום אופן לא קובע

סוגיית ייצוגיות המדגם

בסקרים ארגוניים כל הארגון מקבל לרוב פניה למלא את הסקר, אך אחוז המענה יהיה לרוב פחות מ-100% ולא נדירים אחוזי מענה סביב ה-50%. היות וכך, ניתוח הסקר מתייחס בפועל למדגם ולא לאוכלוסייה כולה

כידוע, ייצוגיות המדגם התלויה בשני פרמטרים שונים לחלוטין:-

- **בגודלו (n)**, הרלוונטי לסוגיית המובהקות הזוכה משום מה לתשומת לב יתירה
- **באקראיות שלו**, כלומר במידה בה אחוז המענה דומה בכל קבוצת עובדים, הן לפי שיוך ארגוני והן לפי שיוך דמוגרפי

בעוד שגודלו של המדגם אינו מהווה לרוב בעיה ברמת כלל החברה ובקבוצות הגדולות בתוכה, סוגיית האקראיות עשויה להיות הרבה יותר משמעותית ובעלת השפעה על מהימנות הממצאים. דוגמא פשוטה עשויה להמחיש את הסוגיה:-

במדגם	באוכלוסייה	
n=50 (25%)	N=100 (16.6%)	מנהלים
n=150 (75%)	N=500 (83.3%)	עובדים

בדוגמא הנ"ל, ברור שקבוצת המנהלים מיוצגת "יתר על המידה" במדגם, ולכן התוצאה הכללית של מדד כלשהו ברמת הארגון עשויה להיות מעוותת עקב כך. חומרנו של העיוות גדלה ככל שקיימת שונות בין הקבוצות השונות

אחת מדרכי הפעולה לטיפול בסוגיה הנה באמצעות שקלול של התוצאות ברמה הכלל ארגונית באופן אשר יעניק את המשקל האמיתי לכל אחת מהקבוצות. מומלץ לבצע שקלול כזה רק במידה וקיים הבדל משמעותי באחוזי המענה בין הקבוצות. בניתוח של כל קבוצה לחוד, נושא האקראיות פחות רלוונטי

סיכום

ניסינו במאמר קצר זה לשפוך אור על נושא קריטי אשר אינו זוכה לדעתנו לתשומת לב הראויה. הדבר עשוי ללמד בכלל על התייחסות בעייתית לכל נושא של סקרי עמדות הן מצד הלקוחות והן מצד הספקים

כאמור, הסיבה המרכזית טמונה בכך שמדובר במערכות בעלות אופי סטטיסטי ולא תהליכי, ולכן קשה מאוד, אם בכלל לאמוד את הנזק הנגרם. דווקא בגלל קושי זה, חשוב להקפיד על גישות מקצועיות מעמיקות ועקביות ולא להתפשר "מאחר וממילא אף אחד לא מרגיש את ההבדל"