

האמנם A1?

קריאה קפדנית בהערכת חלופות של קו הרכבת בין ת"א לירושלים
דן רדר

קו הרכבת בין ת"א לירושלים (וכמובן, גם ההפך) הפך נושא לדיון מקצועי וציבורי מתמשך. אנשי המקצוע מספקים המלצות פעם אחר פעם אך הויכוח על התוואי הרצוי נמשך וההחלטה הסופית (הפוליטית) ממאנת להתקבל. מאמר זה עוסק באופן ביקורתי, בתהליך המקצועי של הערכת החלופות לקו זה משום שכאמור, זהו נושא עכשווי ורלוונטי. קו הרכבת הוא רק דוגמא והלקחים כוחם יפה לכל הערכה עתידית באשר היא בנושאי תחבורה ונושאים תכנוניים אחרים.

בשנים האחרונות נעשו שלוש הערכות לגבי תוואי הרכבת המומלץ בין ירושלים לת"א. הערכה ראשונה נעשתה על ידי צוות בריטי ותוצאותיה התפרסמו ב-1994. הערכה שניה נעשתה על ידי משרד פרטי (חסון-ירושלמי. [אם יש מקום לפרסם את השם](#)) שקיבל את המטלה במסגרת מיקור חוץ מרשות הנמלים והרכבות ותוצאותיה התפרסמו ביולי 1997. הערכה שלישית נעשתה שוב על ידי משרד פרטי (מתת) ותוצאותיה התפרסמו באוגוסט 2000. ההערכה האחרונה כללה רק שלושה מבין שמונת התוואים שנידונו במסמך של 1997 והכותבים הצהירו כי לא היתה בידם שהות מספקת לבחון את שלל ההיבטים ועל כן העבודה נעשתה בחופזה והיא מעולם לא התפרסמה כמסמך רשמי. הצוות הבריטי המליץ על תוואי A (A1 לא היתה קיימת ב-1994) ואילו שתי ההערכות המאוחרות המליצו על תוואי A1.

מאמר זה בודק בעין ביקורתית את תהליך ההערכה אשר הוביל להמלצה זו ומתמקד בהערכה שהתפרסמה ביולי 1997. ההתמקדות בהערכה זו נעשית משום שהיא המסכמת והמאוחרת ביותר מבחינה כרונולוגית ומשום שזו הערכה שאליבא ד'כותביה היא כוללנית ומקיפה את העבודות שנעשו עד אז (נכון ל-1997). הערכה זו פורסמה ביולי 1997 בחסות רשות הנמלים והרכבות בחוברת שכותרתה "המסילה לירושלים ניתוח חלופות ותוואים" ועל כן היא מבטאת עמדה רשמית.

רקע תאורטי

המעריכים בחרו להשתמש בשיטת הערכה רב מטרית אשר ככלל מציבה על ציר אחד של מטריצה חלופות ועל ציר שני מטרות ותתי מטרות שלאורך נבחנות החלופות. כל מטרה ותת מטרה מקבלת משקל המבטא אמירה ערכית לגבי חשיבותה, סך המשקלים 100%. כל חלופה מקבלת על ידי המעריכים ציון לכל מטרה ומטרה. על הציונים להינתן על סמך מספר קריטריונים ברורים כך שניתן לכמת באמצעותם את המטרה ולהימנע ממתן ציונים על סמך תחושת בטן. על הקריטריון לחתור להיות כמותי (לא תמיד זה אפשרי) אך לא בהכרח כספי. הקריטריונים מקבלים משקלות אף הם.

לדוגמא: תחת מטרה אפשרית של "אופטימיזציה הנדסית" קריטריון אפשרי הוא למשל: "מהירות מקסימלית", ואז הציון הוא המספר הגולמי המבטא את המהירות בקמ"ש. קריטריון אפשרי אחר הוא "רדיוס סיבוב" ואז הציון הוא ביחידות מרחק. דוגמא נוספת. תחת מטרה

אפשרית של "כדאיות כלכלית" ניתן להציב קריטריון של "מספר הנוסעים" ואז הציון הוא מספר הנוסעים הצפוי. ברור כי ליחידות מרחק, יחידות מהירות ומספר נוסעים אין מכנה משותף ועל כן יש לתקן כל ציון וציון לפי נוסחה פשוטה ולהביאם למכנה משותף. ללא תקנון אין כל ערך למשקלות ומכאן גם להערכה כולה. הציונים המתוקנים של כל חלופה מוכפלים במשקל ונסכמים. ההמלצה ניתנת לחלופה שזוכה בציון המצרפי הגבוה ביותר. כוחה של השיטה בכך שהיא מעניקה למעריכים יכולת ליצוק לתוך ההערכה משתנים שונים ומשונים וכך להתעלות מעל שיקולי העלות והתועלת שנמדדים על פי רב במובנם הצר ביותר. להלן דוגמא סכמתית:

מצורף כקובץ excel בשם "דוגמא סכמתית"

הערות לדוגמא: 1. כמובן שניתן להציב יותר משתי מטרות או שלוש מטרות משנה או קריטריונים. עם זאת, אין להפריז בניפוח המטריצה משום שאז כל מטרה וקריטריון מקבלים בסופו של דבר משקל זניח. 2. יש לשים לב שסך מכפלת המשקלים היא 100%. תוצאה אחרת מעידה על טעות במתן המשקלות. 3. בעת התקנון יש לשים לב "לכיוון". לעיתים דווקא הציון הנמוך הוא הרצוי.

כל המלצה חייבת לעמוד במבחני רגישות. במבחני רגישות מזינים למטריצה משקלים שונים וציונים שונים מאלו שהוזנו בבדיקה הראשונית ובודקים האם סכימת המטריצה עדיין מורה על אותה המלצה. למשל: האם מתן משקל של 15% לאיכות הסביבה במקום 10% משנה את ההמלצה? ומה ההשפעה של תחזית נוסעים קטנה ב- 20%? או ערך אחר לשעת עבודה? מבחני הרגישות מוודאים באילו תנאים ההמלצה תקפה ומהי מידת המובהקות של ההמלצה.

עד כה התאוריה. מה נעשה בפועל כאשר העריכו ב- 1997 את התוואים השונים? ראשית, למעריכים הידד על שבחרו בשיטת ניתוח רב מטרותית ולא הסתפקו בהערכת עלות/תועלת מצומצמת. למבקרים בכח שיטענו כי הערכה רב מטרותית חשופה להטיה ערכית משמעותית יאמר שגם הערכה במונחי עלות/תועלת כספיים גרידא חשופה להטיות ערכיות ולו רק על ידי קביעת היקף המשתנים שקובעים מהי תועלת ומהי עלות. בנוסף, חשוב לציין שהכסף אמנם נמדד ביחידות אחידות אבל ערך הכסף משתנה ממקום למקום ובין מגזרי אוכלוסייה שונים. למי שיש, הכסף שווה מעט ולמי שאין, המעט שווה הרבה וכך שוב שיקולי העלות והתועלת נמצאים ערכיים.

לשיטת הערכה רב מטרותית יש אם כן, יתרונות רבים אולם אין די בכך ויש ליישם את השיטה כהלכתה.

ליקויים בהערכה

עקביות ואחידות

היסוד לכל הערכה הוא בחינת כל החלופות על בסיס מכנה משותף שיטתי ומושגי. בהערכה הנסקרת נבדקו באופן שיטתי רק 7 חלופות מתוך שמונה על גבי המטריצה המסכמת ואילו החלופה שנקראה חלופה G נבחנה רק באופן מילולי והיא כלל אינה מופיעה במטריצה המסכמת.

כותבי המסמך קבעו מראש כי חלופה G נחותה ואינה חלק מתהליך ההערכה. בהעדר בסיס משותף אין להערכה לגבי חלופה זו כל תוקף.
על ההערכה עצמה להיות קוהרנטית. האמנם?
רשימת המטרות (במסמך עצמו הן מכונות קריטריונים ועל כך בהמשך) המופיעה במסמך ואליה מתייחס הטקסט בהציגו את החלופות, מובאת להלן כלשונה:
"התוויה אופקית ואנכית - קריטריונים הנדסיים", "השפעה על הסביבה", "אומדן עלות",
"איחוד תשתיות", "תאום מערכת הדרכים", "תחנות", "זמן נסיעה", "מספר הנוסעים"
כיצד אם כן מופיעות במטריצה המסכמת במקום "תאום מערכת הדרכים" ו- "תחנות" המטרות הבאות: "מעבר דרך מודיעין" ו "מיקום תחנה סופית". פליטת קולמוס או שמא המטריצה הסופית מסכמת נתונים שאינם מופיעים בגוף המסמך?

ליקויים מתודולוגיים

המעריכים לא מיצו את כל כוחה של השיטה הרב מטריתית ובחרו להעמיד לבדיקה שמונה מטרות ללא תתי מטרות אשר מחדדות את המטרות העיקריות. על פניו אין פסול בכך אולם אי ניצול ההיררכיה שהשיטה מציעה הוביל לכשל שהוא לכאורה סמנטי וכמעט קטנוני אבל הוא מעיד על אי הבנה מהותית של השיטה ומשליך על איכותה של ההערכה ככלל. כפי שנכתב ברקע התאורטי מטריצה של הערכה רב מטריתית מציעה מטרות שלאורן נבחנת כל חלופה וקריטריונים שבאמצעותם מכמתים כל מטרה ומטרה או לפחות מספקים כלי לחידוד הבחינה. במקרה דנן, נראה שהמעריכים לקו בבלבול מושגי וכלל לא הפרידו בין המטרות והקריטריונים. המטריצה המסכמת היא בת היררכיה של דרג אחד המכונה "קריטריונים", ברם, עיון ברשימת "הקריטריונים" מלמד כי מחד, היא אינה מתארת מטרות או יעדים ומאידך היא גם אינה רשימה של קריטריונים. דוגמאות:

"קריטריון הנדסי". אילו היתה הכוונה למטרה אזי הניסוח צריך היה להיות למשל: אופטימיזציה הנדסית והקריטריון צריך היה להיות למשל: זמינות הטכנולוגיה, אמינות הטכנולוגיה, מהירות הנסיעה המרבית, מהירות הנסיעה המרבית בסיבובים וכן הלאה.
"אומדן עלות". המטרה צריכה היתה להיות כמובן מינימום עלות והקריטריון הוא כסף או שיעור תשואה פנימי (שת"פ – IRR).

"זמן הנסיעה". המטרה היא מינימום זמן נסיעה. קריטריון אפשרי הוא זמן הנסיעה ברכבת ובנוסף זמן נסיעה לפי עקרון "מדלת לדלת" (למשל לבניין מסוים במע"ר הירושלמי) שהוא המדד המודרני והוא משקף את ההבנה שכל נסיעה נגמרת ביעד מסוים ולא בתחנה מרכזית. מדד זה משקלל לתוכו את מיקום התחנות יחסית לשימושי הקרקע ומרכזי הפעילות בעיר. (הערה: כל החלופות חולקות אותן תחנות בת"א אך מציעות מספר אפשרויות לגבי מיקום התחנות בירושלים)

העדר קריטריונים ברורים מוליד את הליקוי הבא

חטא אי השקיפות

הערכה רב מטריתית כהלכתה היא שיטתית ומספרית. לא כך במקרה הנידון. מאחר ואין קריטריונים ברורים שמסייעים לכמת את המטרות (בשפת המסמך המטרות הם הקריטריונים) לא ברור כיצד הטקסט המילולי שהוא לב ליבו של המסמך הומר למספרים החד משמעיים כפי

שהם מופיעים במטריצה המסכמת. כיצד פסקה בעברית הופכת ל "6" או "7" במטריצה המסכמת?

אמנם חלק מהטקסט מלווה בתתי מטריצות שעושות הערכת ביניים אולם כיצד לדוגמא תחת המטרה הסביבתית ניתנו ציונים לכל 7 החלופות במטריצה המסכמת, בשעה שתת המטריצה מעריכה רק ארבע חלופות?

מעבר להעדר קריטריונים ומתן ציונים, שקיפות מחייבת זוטות כגון - תיעוד לגבי תהליך קביעת המטרות וקביעת המשקלות שלהם.

כפי שנכתב, מבחני הרגישות הם "בדיקת המעבדה" של כל הערכה ותפקידם לקבוע עד כמה ההמלצה הנובעת מההערכה היא מובהקת ולאילו תנאים היא רגישה.

לא כך בהערכת קו הרכבת בין ת"א לירושלים שכנראה אינו רגיש לכלום ולכן לא נעשו עבורו מבדקי רגישות. הכשל הוא כפול ראשית משום שלא נערכה בדיקת רגישות למשקלות של "הקריטריונים" ושנית משום שלא נערכה בדיקת רגישויות לערכים שהוזנו למטריצה. על מנת לחדד את חשיבות מבחני הרגישות - ב-BART של סן פרנסיסקו תנועת הנוסעים בפועל היותה רק 51% מהתחזיות ולגבי ההפסדים אין צורך לפרט. (Hall, P. (1980) "S.F. BART system", in: Great planning disasters, University of California press, Berkeley).

ערכתי בעצמי מבדקי רגישות. העליתי את המטריצה המסכמת על גליון excel ועל בסיס הציונים שכל חלופה קיבלה במסמך הזנתי משקלות אחרים באופן שרירותי אך הגיוני. הממצאים מופיעים בטבלה הבאה:

בדיקת רגישות למשקלות "הקריטריון"			
בדיקת רגישות 2 משקל אפשרי	בדיקת רגישות 1 משקל אפשרי	המשקל המקורי	"הקריטריון" (כפי שמופיע במטריצה המסכמת)
10	5	10	"קריטריון הנדסי"
5	10	10	"השפעה על הסביבה"
10	15	30	"אומדן עלות"
5	5	5	"איחוד תשתיות"
30	25	25	"מספרי נוסעים"
20	20	5	"זמן נסיעה"
15	15	10	"מעבר דרך מודיעין"
5	5	5	"מיקום תחנה סופית"
100	100	100	סה"כ
A1=6.55 B2=6.75	A1=6.60 B2=6.70	A1=7.15 A=6.55	התוצאה

למרבה התדהמה שינוי מזערי במשקלות (ורק במשקלות) מניב המלצה שאינה A1!
 בדיקת רגישות נוספת שערכתי, הפעם לציונים מערערת אף היא על ההמלצה. חלופה A1 היא החלופה בעלת אורך המנהרות הגדול ביותר ועל כן רגישותה התקציבית למחיר ק"מ מנהרה גבוה. במסמך נקבע כי מחיר מטר מנהרה הופחת מ- \$15,000 כפי שהוערך בעבר ל- \$10,000. במידה וההערכה החדשה אינה מדויקת והמחיר הוא אכן \$15,000 הרי שציון העלות של חלופה A1 קטן אף הוא בשיעור דומה. הקטנתי את ציון העלות מ- 8 כפי שהוא מופיעה במטריצה המסכמת ל- 6, ירידה סבירה ביותר בת 25% והתקבלה התוצאה הבאה: חלופה A וחלופה A1 קיבלו שתיהן את הציון 6.55 אם הציון יורד ל 5 אזי חלופה A1 מקבלת את הציון 6.25. מדוע אם כן ל- A1 הבכורה הבלתי מעורערת?

מבדקי הרגישות שערכתי מדגימים כי ההחלטה רגישה לשיקול הערכי כפי שהוא מתבטא בשינוי המשקלות ורגישה גם למשתנה העלות ואולי אף לעוד משתנים. אמירה זו לא מחייבת היפוך של ההמלצה, יתכן שחלופה A1 היא עדיין החלופה המועדפת אבל חובה על עורכי ההערכה לציין כי היא רגישה להשתנות הפרמטרים שמזינים אותה וכי בתנאים אחרים חלופה אחרת היתה זוכה בבכורה!

טעויות בחשבון

הציונים לא עברו תקנון. תקנון הציונים הוא כלי שנועד להעניק לכל הציונים מכנה משותף. ללא המכנה המשותף לא ניתן לסכם ציונים שעוסקים למשל במרחק וציונים שעוסקים למשל בזמן מבלי שהציונים עצמם במכפלת המשקל יעוותו את הקביעה הערכית שנקבעה על ידי המשקלות. על פניו כותבי התוכנית היו מודעים לכך והם קבעו סקלה **השוואתית** של ציונים שנעה מ - 1 ועד 10. מאחר והסקלה היא השוואתית ולא מוחלטת הרי שהחלופה הטובה ביותר בכל קריטריון אמורה לקבל את הציון 10 ואילו החלופה הרעה ביותר את הציון 1, כל שאר החלופות היו צריכות לקבל ציונים בטווח שבין 1 ל - 10. עיון במטריצה המסכמת, מלמד שלא כך הוא ולא כל הציונים נעו על מלוא הסקלה.

ופרפרא לסיים. במטריצה המסכמת, תחת "הקריטריון" ההנדסי, קיבלו חלופה A וחלופה B1 אותו ציון - 0.7. ואולם, עיון קצר בתת המטריצה שעוסקת בשיקול הנדסי מגלה כי שתי החלופות שונות בכל הפרמטרים. כיצד אם כך קיבלו שתייהן אותו ציון? זמן הנסיעה, כולל תחנות ביניים, של חלופה A1 וחלופה M הוא 33 דקות (או 28 דקות ללא התחנות) כיצד אם כך במטריצה המסכמת, תחת המטרה של זמן נסיעה, חלופה A1 קיבלה ציון 7 ואילו חלופה M קיבלה ציון 8? לאור "פשלה" שכזו מי ילין על זוטות כמו עקביות ושקיפות!

סיכום

אפשר ו-A1 היא החלופה המועדפת אולם לא ניתן להגיע למסקנה זו על סמך ההערכה שנעשתה ביולי 1997. יהיו הסיבות לליקויים שתוארו אשר יהיו וכל קורא/ת/תסיק מסקנות. הלקח הוא כי כל בעל מקצוע ראוי יכול ואף חייב לפצח את "הקופסה השחורה" שהמעריכים עיצבו. אין צורך להתמחות בהערכה, די בזוג עיניים קפדניות והיגיון בריא. לקח זה תקף להערכת תוכניות תחבורה ותוכניות אחרות, לקריאה של מודלים המוזנים במקדמים שונים ומשונים, בתחזיות אוכלוסייה ותחזיות אחרות ובפרוגרמות כמותיות. למעריכים יש אינטרס סמוי ליצור "קופסה שחורה" שאינה ניתנת לפיצוח ואז כוחם "הפוליטי", יוקרתם ונחיצותם המקצועית מועצמים. מנגד, האינטרס הציבורי מחייב שקיפות. ברור כי ללא שקיפות של תהליך ההערכה לא ניתן לעקוב אחריו ולעמוד על איכותו. ללא שקיפות לא ניתן לעמוד על קביעות ערכיות ועיוותים או חלילה, מניפולציות שהמעריכים הטמיעו בהערכה.