

## הערכה ומחקר - האמנם חד הם?

ד"ר מירי לוי-רוזליס

גרסה מתקדמת של המאמר פורסמה ב: **מגמות לטי' (3) 303-319**

### הקדמה

לעתים קרובות בעבודתי בשדה כמעריכה אני נתקלת באי ידיעה מהי ההערכה בכלל ובמה היא שונה ממחקר בפרט. על המדפים אנו מרבים לפגוש ב"מחקרי הערכה". כאשר התוצר הוא בהרבה מקרים יצור כלאיים שאינו עבודת מחקר טובה ולא עבודת הערכה ראויה.

במאמר זה ברצוני לדון בסוגית הקשר והשוני שבין הערכה לבין מחקר. הטענה המרכזית שבכוונתי להציג היא שהערכה ומחקר הן שתי דיסציפלינות נפרדות ומובחנות, למרות הדמיון הקיים ביניהן במושגים, בכלים ובחלק מהשיטות. ברצוני לטעון שחוסר ההבחנה הקיים בין הערכה לבין מחקר פוגע בהערכה. הניסיון המתמשך להפעיל על פעולת ההערכה את אותם קני המידה המופעלים על מחקר, יוצר לחץ על המעריכים לוותר על איכויות המיוחדות להערכה, ובכך פוגם באיכותה וביחודיותה.

### הערכה ומחקר

הערכה ומחקר ממלאים פונקציות שונות בשדה המדעי. מעובדה זו נובעים כל שאר ההבדלים.

מהותו של המחקר המדעי נמצאת אומנם בויכוח מתמיד בין פילוסופים שונים הנוטים להדגיש צדדים שונים שלו: הרציונליסטים מול האמפיריציסטים, הפוזיטיביסטים מול קון (1987). אבל ישנם כמה קריטריונים והגדרות מקובלים. מקובל לומר שהמחקר, כמתודה המדעית, מטרתו לחשוף את החוקים הכלליים על פיהם פועל העולם, ולהבינם, להגדיל את גוף הידע המדעי הקיים, על מנת להשיג שליטה טובה יותר בטבע ועל מנת לספק את התשוקה

האנושית לידע. מכאן שבמחקר, הידע שנצבר הוא הוא המטרה. לידע הטהור יש ערך בפני עצמו.

כמו המחקר המדעי, גם מהותה של ההערכה איננה ברורה וחדה, והיא מצויה בתהליך של התפתחות והשתנות מתמדת. ההגדרות המקובלות מדברות על ההערכה ככלי שרת בידיהם של מתעורבים חינוכיים או חברתיים. תפקידה של ההערכה להעריך עד כמה משיג פרויקט מוערך את מטרתו, מה מושג, מה לא מושג ולמה, ומהם תוצרי הלוואי של פעולתו. בכך מסייעת ההערכה בשיפור הקיים, ותורמת להמשך פעולה שוטף ויעיל של פרויקטים חברתיים/קהילתיים/ חינוכיים או תוכניות התערבות<sup>1</sup>. (Cronbach 1963, 1983; House 1980; Joint Committee 1981; Patton 1981; Scriven 1967; Tyler 1950; Stufflebeam 1971; נבו 1989).

העניין בהערכה התעורר בארה"ב יחד עם וכתוצאה של צמיחת תהליך מוסדר ומובנה של קבלת החלטות, שחדר באיטיות לתחום החינוך הפורמלי והלא פורמלי. כחלק מתהליך זה התעוררה דרישה לבדיקה שיטתית של הנעשה. לפני כן ועד שנות החמישים והששים נבעו יוזמות רבות בתחום העשייה החברתית והחינוכית מגישה פילנתרופית גרידא (Renzuli 1975). שיטות מחקר קונבנציונאליות לא סיפקו את יזמי הפרויקטים, ותהליכי הבדיקה שנבנו היו שונים משל מחקר, וגם נקראו בשם שונה - הערכה. (Rossi & Wright 1984; Tyler 1942). במשך הזמן חלה הינתקות מזרמי מחקר מקובלים והתכנסות אל זרם או גישה מובחנים של הערכה (Rossi & Freeman 1982).

הדרישות מהמעריכים ומהערכה עברו שינויים. מבדיקה של מטרות אופרטיביות ומדידות, בשנות החמישים, דרך הדרישה לספק מידע שימושי למקבלי החלטות, ולעצב את ההתערבות עצמה בשנות השבעים, (Scriven 1967; Stufflebeam et al 1974; נבו 1976). באותה

<sup>1</sup> הערכה מעריכה תוכניות חינוכיות, קהילתיות, חברתיות, תוכניות לימודים, פרויקטים, מצומצמים או רחבים, פשוטים או מורכבים, לזמן קצר או בליווי ארוך טווח, כשהיא מתבוננת במבנה, בתהליך, ביחסים, בתוצרים, ברמת פרט או ברמת מערכת. מטעמי נוחות אכנה לאורך המאמר את כל אלה בשם "פרויקט".

תקופה נוסף הצורך להיות יצירתיים על מנת להצליח להגיב לשונות הרבה שמציבים הפרויקטים השונים. (Patton 1981). בשנות השמונים ציפו מהמעריכים לקחת בחשבון את כל מי שיכול להיות מושפע מפעולת הערכה, להבנות מחדש את מושגים ותמונת עולם עבור הקלינטים, ולהעצים אותם באמצעות הערכה. ( Abma 1997; Drewett and Lincoln 1989; Pawson 1996; Werner 1978 ).1997; Guba

השינויים והדרישות באו על חשבון המערכים הברורים והמכוונים של מחקר. המעריך הוא שנדרש לקבוע איזה סוג מידע חשוב, למי הוא חשוב, איך לאסוף אותו, איך לעבד ולפרש אותו, ומה הן המסקנות וההמלצות הנובעות ממנו.

בגלל הבעיות שנמנו לעיל, נמצאים מעריכים בין הפטיש לסדן. מצד אחד מוצגת להם דרישה לתת תשובות לשאלות רבות, מצד שני קיים הקושי של חוסר במסגרות עבודה ברורות או בכלים יעילים מול המציאות המשתנה. ( Eyken, Goulden & Crossley 1995, Guba & Lincoln 1991 & Dempsey 1991, Tucker ). מעריכים גם מציגים פתרונות לאותם קשיים ומציעים שיטות הלקוחות מתחומים שונים על מנת להתגבר עליהם, כמו למשל ( Finne et al 1995, Kazi 1996, Tucker & Dempsey 1991 ).

למרות ההינתקות של הערכה ממחקר, למרות הדרישות השונות ממעריך ומחוקר, עדין רב הבלבול בין שני התחומים האלה.

ההבדל העיקרי בין הערכה לבין מחקר הוא בתחום ההחלה שלהם. הן המחקר והן ההערכה שואפים להגדיל את הבנתנו בתחומים השונים. אלא שסוג ההבנה הוא שונה. המחקר מבקש לנסח את הידע האנושי בחוקים כלליים ולכן ההבנה שהוא שואף אליה היא הבנה של חוק מופשט ומוכלל שניתן להחילו על מספר גדול ככל האפשר של אירועים.

ההערכה מבקשת לצבור ולהעביר כמשוב ידע והבנה לגבי מערכת קונקרטית של פרויקט או תוכנית. ההערכה מנסה להבין את

ההשפעות ההדדיות בין מקסימום אפשרי של משתנים בזמן נתון ובמקום נתון.

אם במחקר הגדלת גוף הידע המדעי הוא המטרה ויש לו ערך בפני עצמו, הרי שבהערכה הידע הוא אמצעי, ויש לו ערך קודם כל, כמשוב לפרויקט המוערך. בשדה של התערבויות חברתיות וחינוכיות, של שינויים קהילתיים ופוליטיים, הופך הליווי של הערכה כמעט להכרח. היכולת להתבונן מהצד, להבין את התהליכים, לארגן ולנתח אותם, ולהחזיר אותם כמשוב מושכל למבצעים, זו תרומה ובקרה חשובות מאד המשפרות באופן משמעותי את היכולת לפעול בגמישות במצבים משתנים.

מתוך שלל הגישות שקיימות היום לגבי הערכה נראה שניתן להגדיר את תחומי ההחלה שלה באמצעות לוח מס' 1.

כל האמור במאמר זה נכון לגבי ארבע האפשרויות כפי שמוצגות בלוח, ולגבי כל קומבינציה אפשרית שלהן. תוצרי ההערכה צריכים להצליח לפתוח את ה'קופסא השחורה' גם כאשר מדובר בהערכת תוצר מסכמת. (Allterman, Carmon & Hill 1984; Mor 1992).

לוח מס' 1: תחום ההחלה של הגישות המקובלות להערכה.

הערכה מסכמת		הערכה מעצבת		
ותיעוד	סיכום	התהליכים	בדיקת	הערכת
ואופני	התהליכים	בפרויקט	הקורים	תהליך
	היישום של הפרויקט.	ובאופני	היישום שלו,	
		תוך מתן	משוב שוטף	
		למפעילים	ולמקבלי	
		החלטות.		
בה	בדיקת המידה	כפי	בדיקת תוצרי ביניים	הערכת
מטרות	הושגו	ממטרות	שנגזרים	תוצר
	הפרויקט.	ומהרציונל	הפרויקט	
		שלו, ומתן משוב	למקבלי	

---

## החלטות ומפעילים.

---

כאשר מדברים על תחולת ההערכה לעומת תחום ההחלה של מחקר, חשוב להבחין בין שני חלקים עיקריים:

א. פרוצדורות הבדיקה אותן מפעילים אנשי ההערכה או החוקרים, על מנת לאסוף נתונים (שאלונים, ראיונות, תצפיות).

ב. ממצאי הבדיקה, המסקנות הנגזרות מהם והשלכותיהם.

הסיבה להבחנה היא השוני העקרוני באפשרויות התחולה הקיים בין שני הגורמים. כאשר מדובר בהערכה רצוי שפרוצדורות הבדיקה תהיינה בעלות תחולה רחבה. כלומר שיטתיות, ניתנות לחזרה, למעקב ולביקורת. אם יש מקום לפרוצדורות סטטיסטיות אזי חשוב השימוש בפרוצדורות סטטיסטיות מקובלות, ואם ניתן (וברוב המקרים לא ניתן) מוטב גם שימוש חלקי בכלים מוכרים או בכלים שניתן להעבירם במערכים נוספים.

השוני המרכזי בין הערכה ובין מחקר הוא כשמדובר בממצאים ובמסקנות הנגזרות מהם. ממצאי מחקר שואפים להיות מנוסחים באופן המוכלל ביותר כך שיהיו קרובים ככל האפשר לניסוח של חוק כללי. כלומר תחום החלתם רחב.

בהערכה הממצאים צריכים להיות תלויי פרויקט קונקרטי ונכונים לאותו פרויקט באותו זמן. המסקנות מתייחסות אל הפרויקט המוערך ללא כל שאיפה לחרוג אל מעבר לו. (אין זה אומר שחריגה כזו היא בלתי אפשרית, לעתים קרובות ישנם ממצאים בעלי ערך מחוץ לפרויקט המוערך, אלא שלא זו המטרה הראשונית של ההערכה). החשיבות של יכולת ההכללה של הממצאים אל מחוץ לפרויקט או על מצבים אחרים היא מועטה. כלומר תחום החלתם צר.

במקרים בהם פועלים המעריכים בפרוצדורות של מחקר, ומבצעים הערכה במערכים מחקריים (מחקרי הערכה), הם מפעילים כלי מחקר מתוקפים לשאלות אופרטיביות ממוקדות. כתוצאה מכך עולה תחום ההחלה של הבדיקה פלאים. אך מבחינת הממצאים קורה דבר פרדוכסלי: ההתמקדות במשתנים מוגדרים, מוכללים ומופשטים מגבירה את תחום החלת הממצאים אך מקטינה את תחום החלת המסקנות עבור הפרויקט הנתון. לשון אחר: גוברת יכולת ההכללה של ממצאים/מבודדים אך קטנה יכולת המעריך לתת משום שלם על כל הממדים החשובים. מה שקרה הוא שהמערך המחקרי נתן תשובות היאות למחקר. תשובות בעלות יכולת הכללה גבוהה שנאספו באמצעות כלי עם יכולת שחזור גבוהה. אלא שאיכות ההערכה נפגמת מכיוון שהמבצעים השונים בפרויקט לא יוכלו לקבל מסקנות שרלוונטיות להמשך עבודתם, ושנוגעות לחלקים שונים, לקהלים שונים ולשאלות שונות בפרויקט. ממצאים חלקיים המורכבים ממשתנים תאורטיים מבודדים בעלי יכולת הכללה גבוהה המנוסחים בשפה מדעית, קשים מאד לתרגום לתוכנית אופרטיבית או לשיפורה. ולכן יעילותם כמשוב לעשייה מוטלת בספק. (משוב כזה, דרוש לעתים, לצרכים שונים, למשל להצדקת פרויקט או לגיוס כספים, אולם לא לצורך העשייה המתמשכת בפרויקט.)

במחקר קיימת אותה בעיה. תחום ההחלה הרחב של ממצאי המחקר בא לעתים קרובות על חשבון ההבנה השלמה של הנושא הנחקר. (צבר בן-יהושע 1990, Tucker & Dempsey 1991, Geertz 1973). תחומי מחקר שמנסים להתגבר על בעיה זו, כמו חקרי אירוע (case studies) או המחקר האיכותני (צבר בן-יהושע 1990) מתפתחים כל הזמן. בשיטות מחקר אלה מותר החוקר על החלה רחבה של ממצאיו לטובת הבנה רחבה יותר של המורכבות הנחקרת. בכך דומות שיטות מחקר אלה להערכת פרויקטים, אולם אינן זהות שכן מטרתן ותפקידן לא השתנה: הגדלת גוף הידע המדעי באמצעות ניסוח חוקים מוכללים. כלומר השאיפה היא לתחום החלה רחב ככל האפשר של ממצאיהם.

טרנר וולס כתבו על העיסוק במחקר המדעי כמניפולציה על מידע, הנשלטת על ידי מתודות ידועות, והמעוגנת בתיאוריה - יוצאת ממנה וחוזרת אליה. (Turner 1986; Wallace 1968). התיאוריה מארגנת קונסטרוקטים, הכללות, הגדרות ופרופוזיציות שמייצגות ראייה או הצגה שיטתית של תופעה על ידי ספציפיקציה של היחסים בין המשתנים, במטרה להסביר ולנבא את התופעה. (Kerlinger 1972). במחקר המדעי התיאוריה היא נקודת ההתייחסות העיקרית של הגיונות המחקר המקובלים: ההגיון הדוקטיבי וההגיון האינדוקטיבי.

על פי ההגיון הדוקטיבי ההשערות קשורות קשר לוגי תקף להנחות התאורטיות המקדימות. ההשערה היא ה'מוסבר' (האקספלננדום) של הפסוק הדוקטיבי. בהשערה אין חדש ואסור שיהיה בה חדש לעומת ה'מסביר' - האקספלננס - ההנחה התאורטית הא-פריורית. שיטה זו לניסוח השערות טובה למחקר הבודק תאוריה. היא מבטיחה את החוקר מפני סטייה ושומרת שמחקריו יהיו בתחום ההחלה של התאוריה ושמסקנות מחקריו יהיו קשורות לתאוריה ממנה יצא. תופעות המתגלות בשדה אינן בתחום ההחלה של ההגיון הדוקטיבי בכלל. על פי הגיון זה השדה הוא רק כר לבדיקה של ההשערות.

ההגיון הדוקטיבי מנוגד להגיון של עבודת ההערכה, מכיוון שהערכה בודקת את השדה על מנת לגלות את המשתנים והגורמים שמשחקים בו תפקיד ואת הקשרים ביניהם, ולא משתמשת בשדה לבדיקת משתנים והשערות שנגזרים באופן תקף מהנחות תאורטיות מקדימות.

ההגיון האינדוקטיבי. על פי הגיון זה משערים השערות לגבי תופעות מוכללות הנצפות בשדה. ההשערות מכוונות לבדוק את ההסתברות שתופעות אלה תחזורנה על עצמן מעבר לזמן ולמקום, מתוך רצון לנסח חוק הסתברותי כללי. על מנת לדעת את ההסתברות עלינו להכיר את התכונה הנבדקת בקבוצת ההתייחסות המסוימת ואת התנאים א-פריורי (למשל שלמטבע שני צדדים והוא חייב ליפול על אחד מהם). (Copi 1961).

תנאים אלה בדרך כלל אינם מתקיימים בהערכה, ומשתי סיבות: א. הכללה אמפירית היא טענה להסתברות סבירה להופעת תופעה. בשביל לטעון להכללות אמפיריות דרוש שלב מקדים של בדיקת התופעה על מנת להכיר את תכונותיה. הערכה נמצאת לפחות בחלקה, בשלב המקדים הזה או עוד לפניו (מנסה לחשוף מה הן התופעות המשמעותיות ולכן עדיין ללא יכולת לבדוק את תכונותיהן והסתברותן). ב. מטרת ההסתברות היא להגיע להכללה מעבר למקרים פרטיים, לענות לשאיפה לניסוח חוק כללי. בעבור ההערכה, מסיבות שנמנו קודם לכן זו איננה מטרה ראשית.

הערכה אינה תלויה תיאוריה אלא תלויה שדה. ישנן תיאוריות של הערכה ומודלים של הערכה וגישות שונות לשאלות הערכה. ישנן תיאוריות שבאמצעותן אנחנו מפרשים את ממצאי ההערכה. אבל ההערכה במהותה היא תלויה שדה, במובן זה שהשאלות שמציגה פעולת הערכה מסוימת אינן נובעות מתיאוריה, אלא מהשדה, מהפרויקט המוערך ומצרכי המתעורבים ובעלי העניין לרמותיהם השונות.

עבודת המעריך דומה במובן זה לעבודת הרופא או הפסיכולוג. כפי שהרופא או הפסיכולוג נסמכים על המחקר ועל התאוריה בתחומם, מסתמך המעריך על ידע תאורטי רחב בנושאים השונים שבהם עוסקת ההערכה. כפי שטיפולו של הרופא בפצינט אינו בודק תאוריה (להבדיל מהמחקר הרפואי) למרות שהוא נסמך על ידע כזה, גם המעריך, הנסמך על ידע תאורטי, אינו בודק תאוריה.

הנסיון לשמור את ההשערות בתוך תחום תאורטי כלשהו ימנע מן המעריך לשער השערות שאינן בתחום ההחלה של התאוריה שבתוכה הוא עובד. בחירת תאוריה מתאימה להשערות ולמשתני הפרויקט יכולה להעשות בשלב מאוחר יותר, כאשר אנו מחפשים הסבר לממצאינו או לחלקם.



ההערכה מונחית במידה רבה על ידי העיקרון של אבדוקציה ( Peirce, 1979) עקרון לפיו נגשים לבדיקת השדה ללא הנחות יסוד מקדימות, ללא השערות וללא

תאוריה<sup>2</sup> כל אירוע נבחן ונבדקת חשיבותו. השערות מועלות ביחס אליו: האם הוא קשור לאירועים אחרים וכיצד, או שמא מדובר באירוע מבודד, ומה משמעותו.

לפי עקרון זה, המעריך מגיע ללא השערות מחקר ומושגים מופשטים הנגזרים מתאוריה הממקדים את התבוננותו. ישנן כמובן מטרות ההתערבות, ומושגי ההתערבות, והמעריך הוא איש מקצוע, בעל ידע. אין הכוונה לטעון שהוא פועל אינטואיטיבית או ללא ידע מוקדם, אלא שהוא פועל, ככל האפשר, ללא הנחיות תאורטיות לטיב הממצאים שעליו לחפש ולמצוא. במידה מסוימת דומה הדבר לעבודתו של הבלש המדגיש את הצורך להיות נקי מהשערות ופתוח לכל אפשרות ומידע שמציע להם הנושא הנחקר. ההנחייה למעריך, כמו לבלש היא: בחן את כל זירת ההתרחשות מבחוץ ומבפנים, שאל את עצמך מה ראית, לא מה ציפית לראות, או מה קיווית לראות, אלא מה ראית. (James 1989).

אין הכוונה לטעון שהמעריך בודק את כל העובדות שבשטח, בעוד המדען בודק רק את אותן עובדות שרלוונטיות לשאלת המחקר שלו. בדיקה כזו איננה אפשרית, וגם אין בה כל צורך. אלא שבמדע מהווה התאוריה מסגרת התייחסות שמכתיבה אילו נתונים הם רלוונטיים ואילו אינם, ובהערכה מגדירים את השאלות והמשתנים - צרכי הפרויקט והמציאות שבה הוא פועל. אנשי הפרויקט, יחד עם המעריכים מגדירים בעיות. אנשי ההערכה בודקים אותן בשדה. השאלות עוסקות ביחסים בין המשתנים הפרטניים הרלוונטיים בפרויקט.

---

<sup>2</sup> אין הכוונה ל-Grounded Theory (Glaser & Strauss, 1967) למרות הדמיון לכאורה. שכן גישה זו מדברת על יצירת תיאוריה, זו מטרתה החשובה. השוני הוא רק בכיוון ובשיטה. בהערכה אין בניית התיאוריה מטרה כלל ועיקר.

כך למשל ידובר בגילן של עשר המדריכות בפרויקט "צעד בצעד" ולא 'גילי' בכלל, השכלת אותן עשר מדריכות בפרויקט ולא 'השכלה' בכלל, והשאלה לגבי ההתמודדות של אותן עשר מדריכות בפרויקט "צעד בצעד", שטווח השכלתן נע בין 10 ל- 12 שנות לימוד, עם מורכבותו של שלב ב' בתוכנית החדשה של הפרויקט. למדריכות יש שם, ולתוכנית יש שם. לא מדובר פה בשאלה מופשטת ומוכללת של קשר בין רמת השכלה, מורכבות מטלה ואיכות הביצוע שלה.

בהרבה מקרים מדובר במדגם קטן שהופך את הבדיקות הסטטיסטיות לקשות מאד או בלתי אפשריות. השונות שגוברת עם הקטנת המדגם, ושמהווה עבור החוקר הפרעה, היא חלק מתחומי התבוננותו של המעריך.

ההערכה אינה שואפת לנפות משתנים רלוונטיים מתוך הסיטואציה, משתנים ותופעות שאינם רלוונטיים לצורך בדיקת ההסתברות של חוק כללי רלוונטיים מאד לעבודת המעריך. למשל תופעות מיוחדות לפרויקט כמו אופיים המורכב של אנשים שונים בפרויקט, או גודל או צורה של מבנה, טיב השכונה שבה נעשית הפעילות (ולא בהשוואה עם מדגם מייצג של תופעות דומות).

ההשערות נבדקות באופן שיטתי, באמצעות מירב הכלים האפשריים (ראיונות, תצפיות, שאלונים, ניתוח מסמכים) - אופרציונליזם מרובה - ורק לאחר שממצאים שונים מתכנסים לעבר אותה מסקנה, ניתן לומר שיש לנו ממצא בעל משמעות.

לדוגמא, בפרויקט שמפעיל מספר תוכניות היתה בעיה של חוסר יוזמה אצל המפעילים הסמך-מקצועיים של התוכניות. מספר השערות צצות מיד: התפקיד אינו ברור ולכן לא אפשרית יוזמה, הצוות אינו מתאים לתפקיד (כולם?); הוראות ההפעלה נוקשות ומחמירות וחריגות זוכות ל"ענישה"; יש ניהול צנטרליסטי הקוטע יוזמות; קשה לשנות דברים ונוצר חוסר אונים נלמד; מדובר בדפוס תרבותי של האוכלוסייה בה פועל הפרויקט (וממנה באים

המפעילים); המפעילים לא קבלו הכשרה מספקת, הם חסרי בטחון ולכן חוששים לחרוג.

תוך כדי המשך ההכרות עם השדה מופרכות חלק מההשערות (מסתבר שהמפעילים מקבלים הכשרה שבועית רחבה, אישית וקבוצתית ושהאווירה היא תומכת ואין ענישה על חריגות, והניהול הוא דמוקרטי עם אוטונומיה רבה לעובדים). מתחזקות השערות אחרות, ובעיקר מתחזקת ההשערה שמקור הבעיה הוא בדפוס התנהגות תרבותי. מכאן אפשר ללכת הלאה לבדוק את ההשערה המאוששת בחלקה, באופן יותר מובנה, לבדוק אם היא קיימת גם בחלקים אחרים של הפרויקט, ולחפש פתרונות שיעלו בקנה אחד עם מקור הבעיה, פתרונות שיהיו מעוגנים תרבותית.

את המציאות חושף המעריך, באמצעים המדעיים ביותר העומדים לרשותו, באופן השיטתי ביותר, באמצעים הניתנים ככל האפשר לשיחזור באותו פרויקט. כאשר השדה ותנאיו מכתיבים את כלי הבדיקה (התאמת השפה ואופי הכלי: אין טעם להעביר שאלון לאוכלוסייה שמתקשה בקריאה), ואת מושגי הבדיקה (מושגים ושפה של הפרויקט ולא של הז'רגון המדעי). בכך דומה פעולת ההערכה למחקר האיכותני והפנומנולוגי, אשר, כפי שנאמר, מטרתיהם שונות.

המחקר האיכותני והפנומנולוגי שואפים להגביר את הבנתנו בהתנהגות האנושית, ולכן למצוא משתנים משמעותיים הניתנים להכללה ולניסוח בצורת חוק. לצורך זה מנופים בין תהליך הבדיקה של התופעה להסבר שלה משתנים הקשורים לאירוע הספציפי שאינם ניתנים להכללה או שאינם קשורים למערך ההסברי המוצע. פעולת ההערכה משתדלת להשתמש בכל התופעות שחשפה לצורך מערך הסברי קוהרנטי של הפרויקט הנבדק, ובעצם מקטינה בכך את יכולת הכללה של ממצאיה.

בתפישת הקוהרנטיות דומה פעולת ההערכה למחקר האתנוגרפי ההוליסטי. אלא שהמחקר האתנוגרפי, שנערך בסביבה טבעית, עוסק

במשתנים וקטגוריות שמוטבעים בתפישה האנתרופולוגית: מוסדות וארגון חברתי, מבנה שארות, פעולות חברתיות, טכסים, סמלים ותוצרי תרבות, ונערך גם הוא מתוך מטרה להגיע להכללות אודות ההתנהגות האנושית ולהוסיף לידע האנושי. (אשכנזי 1986).

המעריך מארגן את ממצאיו בתבניות ברורות משמעות ואם אפשר ניתנות להשוואה. לעתים ניתן להכליל מהממצאים הפרטיקולריים למבנים ותהליכים מוכללים יותר, או להגיע לתיאוריה מסבירה. אולם המטרה העיקרית היא למצוא הסבר קוהרנטי למה שקורה בפרויקט, הצלחותיו וכשלונותיו.

לאחר שמצאנו בפרויקט מסוים שמנהלו הסמכותי מדי קוטע כל יוזמה אצל עובדיו, אפשר לצאת להסבר מוכלל ונתמך תאורטית, על הקשר שבין סמכותיות ליוזמה. לא ניתן ללכת הפוך ולצאת מהשערה על קשר כזה מכיוון שאין לנו שום דרך לדעת מראש איזה סוג מנהל, דפוס ניהול או אינטראקציה עם עובדים תמצא בפרויקט, או אם שאלת האינטראקציה בין המנהל לעובדיו תעלה בכלל על הפרק. (וכן, למעריך צריך להיות ידע מקדים מספיק על דפוסי ניהול, על סמכותיות, על יוזמה, או שעליו להעזר במי שמקצועו בכך).

פעולת ההערכה אינה נובעת מתאוריה, אבל הפרויקט עצמו הוא במקרים רבים יישום של תאוריה. התוכניות ושיטת הפעולה נגזרות מרציונל תאורטי מסוים או מהשקפת עולם של יזמי הפרויקט, מפעיליו או מממניו. נושא זה של טיב הרציונל שמאחורי הפעילות, הקשר בינו לבין הפעילות בשטח, והאם התוצאות עולות בקנה אחד עם ציפיות הרציונל, או התאוריה על פיה פועל הפרויקט, נבדק על ידי המעריכים כחלק מהבדיקה המקיפה של הפרויקט. בנוסף, אם הפרויקט נובע מתאוריה, השאלות שנבדקות על ידי ההערכה נבדקות בדיאלוג עם התאוריה.

המתודולוגיה המחקרית נובעת מהתיאוריה, נובעת מההשערות הנגזרות מהתיאוריה או מהכללות אמפיריות. היא פרי תכנון מוקדם וקפדני של תהליך האופרציונליזציה והאינסטרומנטליזציה של

ההשערות - תהליך של התאמת הגדרות אמפיריות, הניתנות למדידה - להשערות. כלומר ניסוח במונחים תצפיתיים. מונח תצפיתי צריך לעגן במציאות: לאתר מצב מתאים לעריכת התצפית המתבקשת או ליצור באופן מלאכותי מצב הדומה לזה ככל האפשר. לאחר מכן נערכות התצפיות המתאימות הבודקות את ההשערה. לצורך בדיקה זו בוחר החוקר את המשתנים הנראים לו מתאימים ביותר למושגי ההשערה, את הנבדקים המתאימים ביותר, את השדה המתאים, המערך המתאים.

המתודה של ההערכה שונה מהמתודולוגיה המחקרית. אם החוקר יכול להחליט על נושא חקירתו ועל אופן החקירה - ומבחינה זו ניתן לכנותו אקטיבי - הרי שהמערך הוא ריאקטיבי ולעיתים פרואקטיבי. הוא מגיב למאורעות ולעיתים מנסה לכוון אותם. אין באפשרות ההערכה לבחור לעצמה נבדקים, מערך או משתני מחקר. השדה נתון, האוכלוסייה נתונה ומשתני המחקר, לפחות בחלקם, אינם ידועים מראש. ישנה הגדרה כללית של שאלות ההערכה. שאלות ההערכה אינן מנוסחות כהשערות והמשתנים אינם אופרציונליים. לעיתים אפילו ההגדרה הנומינלית של מושגים בפרויקט היא שאלה להערכה. כלי ההערכה (ראיון, שיחה, תצפית, שאלון, צילום בוידאו, ניתוח פרוטוקולים, ניתוח דיאלוגים או כל כלי אחר), מתוכננים ונבחרים על פי טיב האוכלוסייה, הפעילות, שאלות ההערכה, הזמן והתקציב וגם החוזה<sup>3</sup> שיש למעריך עם מפעילי הפרויקט או עם יוזמי ההערכה.

המעריך אינו בודק משתנים מבודדים אלא אירועים. בדיקה של אירוע כוללת בתוכה את מרב המשתנים האפשריים עם הקשרים ההדדיים ביניהם, וגם גורמים שאינם משתנים (סוג השכונה, אופי המנהל). מתוך ניתוח האירועים מגדיר המעריך קטגוריות ומשתנים.

<sup>3</sup> החוזה - זהו עניין חשוב בהערכה. ישנו כמובן החוזה הפורמאלי. אבל ישנו החוזה שבין המעריך לאנשים בפרויקט. בדרך כלל לא חוזה כתוב. לעיתים גם לא מנוסח בעל פה. מדובר ב"הבנות" שיש למעריך עם המוערכים על גבולות החדירה שלו. על אופני המעורבות שלו (האם מותר לו או אסור לו להיות נוכח בהנחיה אישית של מדריכה סמך-מקצועית? האם ישיבות הנהלה פתוחות או סגורות בפניו? האם הוא מביא את ממצאיו ומסקנותיו בפני כל הצוות? דק הנהלה? בכתב או בע"פ?).

ההגדרה יכולה להעשות על סמך ידע מקדים, על פי מטרות הפרויקט, או על פי קריטריונים אחרים. למשל, הופעה של משתנה שחוזר באופנים שונים במקומות שונים בפרויקט, כמו עמימות בהגדרות של תפקידים שונים בפרויקט שחוזרת ברמות שונות (ובמקומות שונים). המשתנה נבדק בהקשרים הרלוונטיים לפרויקט, על מה הוא משפיע, מה הוא משרת, ממה הוא נובע. האם בכל הפעמים שמשתנה זה מופיע, הגורם הוא אותו גורם? ההשפעה היא אותה השפעה?

כדי לאסוף נתונים משתמשת ההערכה בכל השיטות המקובלות במדעי החברה ובמדעי הרוח. בדרך כלל משתמשים בהערכה ביותר מאשר כלי אחד או שיטה אחת לאיסוף נתונים. ההערכה עוסקת בתופעות רבות: יחסים בינאישיים, הלכי רוח, סגנונות ניהול, ערכים, אמרות כנף, סגנון לבוש. ריבוי השיטות והכלים חשוב לצורך יצירת אופרציונליזם מרובה. אופרציונליזם מרובה ואישוש הממצאים באופני בדיקה שונים (תגובתיים כמו ראיון ולא תגובתיים כמו ניתוח מסמכים), הוא אחד מכלי העבודה החשובים של המעריך וקריטריון חשוב לממצאיו.

המעריך צריך להיות בעל יכולת מקצועית גבוהה שתאפשר לו גמישות בשיטות איסוף נתונים, יכולת לבנות כלי מחקר המתאימים לקהל היעד ולנושא הנחקר. שכן, כמו במחקר האיכותני או האתנוגראפי, המעריך עצמו הוא כלי, שבנוסף לכישוריו האקדמיים והמקצועיים ישנה חשיבות רבה גם לכישוריו האישיים, ליכולתו ליצור קשר ותקשורת, לרכוש אמון, להעביר את ממצאיו כך שיהיו ניתנים לקבלה וליישום (ולא רק להבנה). למרות חשיבות אישיותו של המעריך בתהליך, הערכה היא עבודה שיטתית, מובנה ומושכלת ולא 'אינטואיציה' או אומנות.

אחת השאלות העולות תדיר היא: מדוע הדגש על הפרויקט הקונקרטי, האם אין כל הפרויקטים דומים בסופו של דבר? פטון מתאר את פוטנציאל השונות הרבה הקיים בין פרויקטים שונים. הוא מגדיר עשרים משתנים, כמו למשל, אופי התוכני(יו)ת, זמינות

הנתונים, מבנה תהליך קבלת ההחלטות, רמת ההשכלה של הצוות, של הצרכנים וכדומה. אם לכל משתנה כזה יש רק שלושה ערכים (שני ערכי הקצה והאמצע), מתקבלים 8,000 צרופים אפשריים. גם אם ישנם צרופים שסבירותם נמוכה, זהו עדיין מגוון רחב מאד. השדה שבו פועל המעריך הוא שדה שמשתנה כל הזמן גם בגלל השונות שבין הפרויקטים וגם בגלל השתנותם ההמשכית. כמעט ולא יתכן שהסבר שלם הלוקח בחשבון את כל המתרחש בפרויקט על רמותיו השונות יהיה נכון לפרויקטים נוספים (Patton 1981).

אם כל פרויקט שונה מרעהו, ודורש התייחסות שונה, עולה השאלה של תוקף ממצאי ההערכה ושל האפשרות להכללתם.

במונחים של תוקף מערך המחקר או מערך ההערכה, מקובל לומר שבעוד שהמחקר והחוקר שמים דגש על תוקף פנימי וחיצוני, מתמקדת ההערכה, בגלל התענינותה בפרויקט המסוים, כאן ועכשיו, בתוקף פנימי בלבד.

הפרקסיס של ההערכה מתגבר בקלות יחסית על רוב האיומים לתוקף פנימי שמקורם במשתני נבדקים (היסטוריה, בשילה, ברירה, נשירה, אינטראקציה עם ברירה, התמרמרות נבדקי קבוצת הביקורת, היסטוריה מקומית). הערכה, בניגוד למחקר עוסקת בריבוי משתנים ולא בבידודם, ולכן משתנים אשר במחקר היו נקראים 'ארטיפקט', בהערכה דווקא הם בעלי עניין.

תוקף המבנה של כלי מדידה משמעו הלימה בין ההגדרה התצפיתית לבין ההגדרה הנומינלית של המשתנה הנחקר. (פרס ויצוב, 1984). בעית התוקף מתעוררת כיוון שמדידה במדעי החברה היא, עם יוצאים מהכלל מועטים ביותר, בלתי ישירה. בנסיבות אלה לעולם אין לחוקרים בטחון מלא שהם מודדים בדיוק את אותה התכונה שהם מתכוונים למדוד (נחמיאס ונחמיאס, 1982).

האיומים על תוקף המבנה (אי בהירות המשתנה התאורטי, חפיפה לא מספקת בין המשתנה התצפיתי לתאורטי), אינם מהווים בעיה

בהערכה מכיוון שההגדרות האופרציונליות והשאלות בהערכה נובעות מהשדה ולא מתאוריה.

זה יותר מעשור נוטים הכותבים להגדיר את התוקף בצורה עמומה ופתוחה לשיפוטים סובייקטיביים. כך למשל קרונבך (Cronbach 1983 p. 108) בטענתו שתוקף הוא סובייקטיבי יותר מאשר אובייקטיבי. קרתוול השווה תוקף לשכנוע, אמינות וקונצנזוס (Krathwohl 1985). קמפבל, אביהם של מושגי התוקף הפנימי והחיצוני, חזר בו ממושגיו אלה כמעט לחלוטין. (Campbell 1986). הוא מציע ניסוח והגדרה מחודשים. את התוקף הפנימי הוא מציע להחליף ב- "Local, Molar, Pragmatic, a-Theoretic) Causal Validity". ניסוח מחודש זה שואל בעצם אם סך כל חבילת הטיפולים הזו, יוצרת שינוי אמיתי, בעשייה היחידאית הזו, במקום המסוים הזה ובזמן המדויק הזה.

לדברי קמפבל קיימים קשרי סיבה ותוצאה הנראים בשטח שמקורם אינו בתיאוריה ולא ברור איך מכלילים אותם. ובכן, יש להפוך את הסדר המקובל שבו תיאוריה קודמת לפעילות. סיבה ותוצאה הנראות בשטח ואינן נובעות מתיאוריה, עשויות להיות מנוף לפיתוח תיאוריה.

התוקף החיצוני שואל על יכולת הכללה, על קשרים קבועים וניתנים לניבוי בין תופעות אמפיריות. במחקר מתכוונים בדרך כלל ליכולת הכללה של האפקט אל מערכים, אוכלוסיות, מצבים וזמנים אחרים, על מנת שניתן יהיה לנסח חוק כללי. שאלה זו חשובה לחוקרים, השואפים לנסח חוק כללי. שאלה במחקרתהיה 'האם התפתחות של יחסים חברתיים בין אנשים החיים יחדיו במשך ימים רבים, 24 שעות ביממה, היא בעלת דפוס דומה מעבר לאוכלוסיות ומצבים שונים. בהערכה נוכל לומר שמעניין אותנו מה קורה בגיבוש של מועמדים לקורס טיס, לצורך הפקת לקחים פיקודיים, או מה קורה במחנה קיץ של בני נוער, לצורך הגברת שביעות רצון המשתתפים, ולא מעניין אותנו אם אותם קשרים יהיו גם בין דיירים של בית אבות, או של פנימייה. לכן בהערכה שאלת התוקף החיצוני, ויכולת ההכללה של הממצאים, חשובה פחות.



לעומת זאת יש חשיבות לשאלת לגבי יכולת להכליל את הטיפול - הפרויקט. בפרויקט היזמים או המנהלים מפעילים טיפול על מנת להשיג אפקט. ההערכה בודקת, במידה ושאלה זו מעניינת, את יכולת ההכללה או היקף ההחלה של הפרויקט. כלומר מדובר ביכולת הכללה של הטיפול המופעל (האם ניתן להעביר אותו אל אוכלוסיות אחרות, מקומות אחרים).

לעיתים, כאשר נעשית הכללה מתוך בדיקה מדגמית של אוכלוסיית הפרויקט אל כלל אוכלוסיית הפרויקט, עולה השאלה של יצוגיות המדגם, היכולת להכליל מהמדגם לכלל אוכלוסיית הפרויקט.

בהערכה אנחנו מעוניינים להכליל מחלק אחד לחלקים אחרים של הפרויקט, ממערך משתנים אחד לאחר, כדי שלא לתאר פרויקט כשורה של אירועים בלתי קשורים. לדוגמא: האם חוסר הגבולות המאפיין את הגדרת התפקידים בצוות הבכיר של פרויקט מאפיין גם הגדרות של תפקידים אחרים באותו פרויקט? האם חוסר ההכרות עם הגדרות של מושגים מרכזיים בפרויקט אופייניים רק לבית ספר אחד (בעיה בית-ספרית) או לכל בתי הספר בפרויקט (חוסר הצלחה של מפעילי הפרויקט להעביר את המסר שלהם). השערות אלו לגבי האפשרות להכליל ממצאים עוזרות למעריך לצייר תמונה קוהרנטית, להבחין בתופעות יוצאות דופן (יעילות על רקע חוסר יעילות או ההפך) ולנסות לזהות את מקורן (חוסר מודעות לחשיבות ההמשגה או חוסר רצון להתמודד עם מורים עקב אי אמונה ביכולתם). ההשערה לגבי האפשרות להכללת ממצא מסוים היא כלי חשוב בידי המעריך להבנת השדה שבו הוא פועל. אישוש השערה כזו נותנת בידיו מידע חשוב שרלוונטי עבורו ועבור הפרויקט שאותו הוא משרת.

מידת הרלוונטיות של עבודת הערכה היא קנה מידה חשוב מאד לאיכותה. חלק חשוב מתפקיד המעריך הוא מתן משוב יעיל למקבלי ההחלטות. מבחינה זו המעריך הוא נותן שרות והשרות צריך להיות יעיל. השאיפה ליעילות המשוב ולתועלת הגלומה בו מדגישה את חשיבות הרלוונטיות שלו. משוב יעיל ורלוונטי תלוי בטיב הבדיקה.

במחקר השאיפה היא להכללה שהיא מעבר לשדה שבו נערך המחקר, ואין שאלה לגבי רלוונטיות המשוב או הממצאים, לשדה בו בוצע. הקריטריון הוא קידום המדע והצטברות הידע.

השאלות במחקר נגזרות מתאוריה, השאלות בהערכה נגזרות מרציונל הפרויקט ממרכיביו ומצרכיו. גם הערכת תוצר מסכמת שאינה מתייחסת לתהליכים, מתייחסת למוצר קונקרטי בפרויקט קונקרטי באוכלוסייה נתונה ובזמן נתון.

השאלות שנבדקות בהערכה צריכות להיות אלה שהתשובות להן עשויות להועיל למפעילי הפרויקט נוכחים או עתידיים, לבצע אותו טוב יותר, למקבלי ההחלטות לבצע את השינויים המתבקשים, ליזמים לעדן את תוכניותיהם. שיקולים אלה אינם עומדים בפני המחקר, ועלולים אפילו לפגוע באמינותו.

יותר מזה, ההחלטה אילו שאלות הן רלוונטיות ואילו לא, במה כדאי להשקיע את משאבי ההערכה ובמה לא, נושאים אלה נתונים במקרים רבים למשא ומתן בין המעריך לבין אנשי הפרויקט. "עניין" בפני עצמו, שהוא סיבה מספקת למחקר, אינו סיבה מספקת לניסוח שאלת הערכה. לעומת זאת, שאלות המחקר אינן עניין למשא ומתן, אלא בין החוקר ועמיתיו וההחלטה על מה לבדוק ואיך, היא החלטה בלעדית של יוזמי המחקר.

הפעלה של פרוצדורת בדיקה מבוקרת, העוסקת במשתנים מוכללים ומופשטים, אינה משרתת כראוי את המעריך ואת מזמיני ההערכה, ואיננה עומדת בקריטריונים של רלוונטיות מכיוון שהתשובה שהיא יכולה לספק לאנשי הפרויקט מתרחקת מעולם התוכן המיידי שלו. בדיקה במערך מחקר מובנה, של מספר משתנים משמעותיים, יכולה להיות חלק ממערך הערכה רחב יותר, המכיל בדיקות נוספות, משתנים נוספים ושאלות על קשרים בין אירועים וההשפעות ההדדיות ביניהם. "משתנים מבודדים" "אוכלוסייה נקיה" "יכולת הכללה" כל אלה לא קיימים בשדה שבו פועלת ההערכה. בפרקטיקה,

המשתנים אינם מבודדים, האוכלוסייה אינה נקיה ויכולת ההכללה היא לעתים קרובות מוגבלת ביותר.

פרויקט נוגע לגורמים שונים שמשפיעים על תפקודו. הוא נמצא בקשר עם אוכלוסיות שונות: יזמים, מממנים, עובדי אדמיניסטרציה, מפעילים ברמות שונות, וישנה אוכלוסיית המטרה שהיא במקרים רבים מגוונת. אוכלוסיות מגוונות משמעו השקפות עולם מגוונות, מטרות שונות ומגוון ערכים שלא תמיד זהים. ישנם בעלי עניין שונים, שלהם דרישות משלהם, ישנה התוכנית המופעלת, שיכולה להיות פשוטה וחד מימדית (שיטה ללימוד עברית לעולים) או מורכבת (שינוי דפוסי חיברות של קהילת עולים מאתיופיה). ואפילו למבנה הפיזי יכולה להיות השפעה (ריחוק או קרבה לאזורי המגורים למשל, מקום סגור או פתוח). כל הגורמים האלה, עם השפעותיהם ההדדיות, הם רק חלק מהגורמים שמשפיעים על האפקט של פרויקט.

רוב פרוצדורות הבדיקה המחקריות לא תתיחסנה לעולם לכל הגורמים האלה יחד ולהשפעות ההדדיות ביניהם, משום שאלה אינן בתחום ההחלה של מחקר. מבחינת המחקר הם מהווים הפרעה. אם נשתמש בלשונו של קון (Kuhn 1962) אלו אינן סוג החידות הנובעות מפרוצדורה של מחקר תקני. פרוצדורות מחקר כמו מערכים דמויי-ניסוי, (quasi-experimental) (Cook & Campbell 1976), המחקר האיכותני, המחקר האתנוגרפי וחקר המקרה, מנסים להתמודד עם הבעיה של ריבוי משתנים. חוקר איכותני עשוי להתעניין בהשפעות ההדדיות בין משתנים מוכללים ומופשטים שונים, אבל הוא איננו מתעניין בהשפעות של משתנים קונקרטיים, בפרויקט נתון, על הצלחתו או אי הצלחתו של אותו הפרויקט. למעריכים, שאלות אלה הן מהות העניין. החוקר בוחר את שאלות המחקר על פי תחום הדעת, עניין וחשיבות. המעריך, על פי הסיכוי לתועלת שבתשובות עבור הפרויקט אותו הוא משרת.

משוב רלוונטי הוא משוב בר שימוש. על מנת שיהיה בר שימוש הוא צריך לעמוד במספר תנאים:

מידיות - המשוב, כדאי שיהיה מידי, או לפחות בתוך פרק זמן שמאפשר את לימוד הממצאים והמסקנות ויצירת או תכנון שינוי. אחרי תקופה ארוכה מדי לא יוכל המשוב להביא תועלת לאותם מקבלי החלטות או מבצעים אליהם כוון. גם כאשר מדובר בהערכה מסכמת, שמתרחשת לאחר מעשה, חשוב שמסקנותיה יגיעו בעיתוי המתאים לתכנונים עתידיים.

פרטיקולריות - המשוב צריך להיות פרטיקולרי כדי שיוכל לתת תשובות העוסקות באופן ישיר בפרויקט בו מדובר. כלומר, לדבר ישירות על המפעילים ועל הצרכנים ועל התוכניות של הפרויקט המדובר, ולא על מושגים מוכללים כלשהם. למשל: ילדי קבוצת "רקפת" שבפרויקט לא נטו לשחק עם ילדי קבוצת "כלנית". הנימוק שנתנו לכך רוב הילדים הוא שילדי "כלנית" הם סנובים. המשך הבדיקה ינסה להבין איזו משמעות נותנים הילדים למושג "סנוביות" ומה מקור התחושה. טענה מוכללת שאומרת למשל ש"פערים מסוג X בין ילדים גורמים לניכור", דורשת בדיקות נוספות, הגדרות אופרטיביות אחרות, ולא תמיד ברור מה התועלת המיידית שלה לפרויקט המסוים. לעומת זאת הממצא הקונקרטי המגלה יריבות היסטורית בין משפחות של ילדים דומיננטיים בשתי הקבוצות, מאפשר טיפול ממוקד ויעיל במתח שביניהן.

שפה ומונחי הפרויקט - כדאי שהמשוב יהיה מנוסח במונחי הפרויקט, או לפחות בשפה המובנת למפעילים וליזמים (גם אם לעיתים אין הגדרותיהם ומונחיהם עולים בקנה אחד עם הז'רגון המדעי). תשובות המנוסחות בצורה של חוקים מוכללים ובשפה מדעית המורכבת ממושגים מופשטים היא במקרים רבים חסרת תועלת למקבלי החלטות או למבצעים בשטח, מה עוד שאינה תמיד מובנת למי שאינם אנשי מדע. מסקנה כמו: "בתוכנית 'כוח לבריאות' של פרויקט 'קהילה וחברה' מצליחות המדריכות האתיופיות יותר מהמדריכות שאינן יוצאות אתיופיה, ללמד נשים אתיופיות להשתמש באמצעי מניעה, מכיוון שהן דוברות את שפת הנשים ומבינות את הקודים התרבותיים שלהן" עדיפה על "סוכני שינוי מקרב הקהילה המטופלת יעילים יותר בהחדרת שינוי בהרגלי

האישות". דבר זה נכון לגבי כל הגישות להערכה, מסכמות ומעצבות, הערכות תוצר והערכות תהליך.

שיפוט ערכי: לא ניתן לדבר על רלוונטיות המשוב מבלי להתייחס לאחד הגורמים החשובים להבחנה בין הערכה לבין מחקר, והקשור לרלוונטיות הממצאים: החלק ההערכתי-שפיטתי של ההערכה. ההיסטוריה של ההערכה מראה לנו, שמהות צמיחת הדיסציפלינה הזו בצורך שהיה (ועדין) קיים אצל המתערכים לפירוש ולשיפוט של הממצאים. זהו גם מקור שמה - הערכה. הועדה המשותפת (Joint Committee) שישבה על מדוכת ההערכה בסוף שנות השבעים הדגישה בהגדרתה את מהות ההערכה את חשיבות ההחלטה על הערך או השווי של האובייקט המוערך. (Joint Committee 1981).

כמעט בכל אופי של פעולת הערכה, יש צורך בהפעלת יכולת השיפוט של המעריך. אפילו בהערכה מסכמת המתמקדת בהערכת תוצר בלבד (מידת השגתן של מטרות), יש לעיתים קרובות צורך בשיפוט האם המידה שבה הושגה המטרה משמעותית או לא, (Mor 1992), על אחת כמה וכמה בגישות המעצבות הנפוצות היום, שפותחות לדיון את ערכן של המטרות, את ערכן של הפרויקט ואת משמעותם של הנזקים או ההפסדים שעלולים להיגרם בעטיו.

המחקר שואף לאובייקטיביות. הממצא הוא המסר. הערכה, בהיותה כלי עזר למקבלי החלטות, אינה יכולה וגם אינה צריכה להימנע מהמלצות, גם אם הן ערכיות. תהליך הערכה הכולל איסוף מידע בלבד, ללא התייחסות למשמעותו, ללא המלצות, או לפחות הצגת האופציות השונות ועזרה בבחירה ביניהן נראה לכותבים רבים בלתי אפשרי ובלתי רצוי (לוי-רוזליס 1987, נבו 1992,

Scriven 1967, Stake 1969, Joint Committee 1981, Guba & Lincoln 1986, Cordray 1986).

### **סוגיית הסיבתיות**

מימד משמעותי נוסף ברלוונטיות של משוב, הוא הנסיון להבין מה גרם למה, מה נובע ממה, מה סיבה ומה מסובב. רק עם ידיעת הסיבה ניתן יהיה אולי לשלוט בתוצאה. תשובות על קשר בין

משתנים, בדרך כלל אינן מספיקות. נחמיאס ונחמיאס (1982) טוענים, שקיימת דילמה בסיסית בין סיבתיות להכללה: "כדי להבטיח הוכחות בלתי מעורפלות לסיבתיות, מקריבים תכופות את כושר ההכללה." (עמ' 50) כך במחקר. בהערכה יכולת הכללה אינה בתחום ההחלה שלה. אבל הצורך להבין מהיכן נובעות התופעות - בהחלט כן.

לדוגמא, בפרויקט מסוים שבדקנו (לוי-רוזליס ובר-און 1994) היו תלוונות קשות של הילדים על האוכל. "אוכל של מורים" הם כינו אותו ואמרו שהם רוצים אוכל של בית. נסיונות חוזרים ונשנים להשביח את המזון, לא הועילו. במקרה הזה תשובה על קשר היתה נותנת מידע דל ביותר לאנשי הפרויקט ("סוג האוכל וכמות תלוונות של ילדים נמצאים במתאם גבוה ביניהם"). בדיקה מעמיקה של הגורמים לתלוונות העלתה שלגופו של עניין לא חשבו שהאוכל כל כך גרוע, אלא שהאוכל היה אפיק לגיטימי להביע תחושות של אי שביעות רצון, בעוד אפיקים אחרים היו חסומים. מכיוון שהפרויקט היה "פרויקט איכותי", ש"כבוד להשתייך אליו", לא העזו הילדים להביע חוסר נחת. וכך ניתן היה להגיע בכלל להבנה שמתחת לפני השטח קיימת אצל הילדים אי שביעות רצון, והיא נחשפה וטופלה בהצלחה על ידי אנשי הפרויקט. משנפתרה הבעיה נעלמו גם התלוונות על האוכל.

מור (Mor 1992), טוען שגם בהערכה מסכמת - הערכה הבאה לאחר מעשה - לא ניתן להסתפק בתוצאות, אלא חייבים לבחון גם מה נעשה ואיך ולמה הממצאים הם כאלה וממה הם נובעים. אין זה אומר כמובן, שההסבר הסיבתי זמין תמיד, או פשוט לגילוי, או שאנשי ההערכה מתאמצים תמיד לחשוף אותו עד תומו.

ההערכה איננה מוכרת כדיסציפלינה עצמאית, וההתייחסות אליה במקרים רבים היא כאל 'מחקר לא טוב' או 'לא שלם' שאיננו עומד בקריטריונים. נוצר כאן מעגל קסמים: בגלל חוסר הכשר מחקרי ומדעי להערכה כמתודה אקדמית העומדת בפני עצמה, מנסים המעריכים לעמוד בקריטריונים של מחקר, להגיע לציוני תוקף גבוהים יותר משמאפשרת זאת פרדיגמת ההערכה, להגיע לממצאים מוכללים, לבדיקות באמצעות כלי מחקר מתוקפים. ניסיונות אלו גורמים להגבלת מערכי ההערכה ולבדיקת אפקטים ותהליכים מבודדים הנובעים בדרך כלל מהמטרות האופרטיביות שנקבעות על ידי מפעילי הפרויקט או יזמיו.

ההערכה כדיסציפלינה צמחה מלכתחילה מכיוון שלמחקר לא היו תשובות משביעות רצון לסוג השאלות שההערכה מתמודדת איתן. לכן הלחץ על ההערכה ועל המעריכים לעשות פעולות שעומדות בקריטריונים של מחקר, גורם לכך שהמעריכים מתקשים לבצע עבודת הערכה משביעת רצון.

ניסיתי בעבודה זו לעמוד על ההבדלים העקריים, לדעתי, שבין הערכה לסוגיה השונים ובין מחקר לסוגיו. ניסיתי גם להראות היכן קריטריונים של מחקר אינם מתאימים להערכה ופוגמים באיכותה. הכרה בהערכה כדיסציפלינה עצמאית ושונה ממחקר, תאפשר לאנשי ההערכה לפתח קריטריונים ודרכי עבודה שמתאימים לתפקידם.

טבלה מסכמת של הטענות הבסיסיות לגבי הדמיון והשוני בין הערכה למחקר

נושא	דומה	שונה
החלה		מחקר: הגדלת גוף ידע מדעי
החלת בדיקה	רחבה ככל האפשר	הערכה: מתן משוב קונקרטי
ממצאים	החלת	מחקר: רחבה, ממצאים מוכללים.
		הערכה: מצומצמת. ממצאים ממוקדים בפרויקט
תאוריה	ידע תאורטי, מושגים	מחקר: תלוי תאוריה הערכה: תלוית שדה
מתודולוגיה	כלי מחקר	מחקר: נובעת מתאוריה, חוקר אקטיבי.
		הערכה: נובעת משדה, מעריך מגיב.
הכללה		מחקר: שאיפה לניסוח חוק כללי, תוקף חיצוני ופנימי.
		הערכה: דגש על תוקף פנימי והסברים סיבתיים.
רלוונטיות		מחקר: לקידום המדע הערכה: שימושיות לפרויקט.

רשימת מקורות



אשכנזי מ' (1986), מחקר אתנוגרפי באנתרופולוגיה. בתוך: ר. בייט-מרום (עורכת), שיטות מחקר במדעי החברה. יחידה 4. הוצאת האוניברסיטה הפתוחה. רמת אביב.

בלכר, ז' (1987). פילוסופיה של המדע. אוניברסיטה משודרת, משרד הבטחון ההוצאה לאור.

בייט-מרום, ר' (1986) שיטות מחקר במדעי החברה, עקרונות המחקר וסגנונותיו. יחידה 1, עמ' 21.

לויין-רוזליס, מ' (1987). תהליך לבדיקה חלקית של שינוי הנוצר בעקבות התערבות בקהילה. עבודה לשם קבלת תואר מוסמך במדעי החברה והרוח.

לויין-רוזליס, מ' ובר-און נ' (1994) הערכה ואיבחון בכפר הנוער 'אשל-הנשיא' דו"ח פנימי, לא פורסם, הוגש להנהלת כפר הנוער, הנהלת הפנימייה, הנהלת בית-הספר וצוות המדריכים.

נבו, ד' (1976). ההערכה בפרויקט המערך המסייע. בית הספר לחינוך. אוניברסיטת תל-אביב.

נבו, ד' (1982). הערכה בחינוך מהי? היחידה למחקרי הערכה ומדידה. בית הספר לחינוך, אוניברסיטת תל-אביב.

נבו. ד' (1989). הערכה המביאה תועלת. הוצאת מסדה, תל-אביב.

נחמיאס, ד' ונחמיאס, ח' (1982). שיטות מחקר במדעי החברה. הוצ' עם-עובד, תל-אביב.

פרס, י' ויצוב, ג' (1984). מבוא לשיטות מחקר במדעי ההתנהגות, הוצאת אקדמון, ירושלים.

צבר בן-יהושע, נ' (1990). המחקר האיכותי בהוראה ובלמידה. הוצ' מסדה, תל-אביב.

Abma, T.A., (1997), Playing with/in plurality. Revitalizing realities and relationships in Rotterdam. Evaluation. 3(1) 25-48. Sage Publication. London/

- Alkin, M.C., (1968). Toward an evaluation model: a system approach, CSE Working Paper No.4 Center for the Study of Evaluation, Los Angeles
- Alterman, R., Carmon, N., Hill, M. (1984). "Integrated evaluation: a synthesis of approaches to the evaluation of broad-aim social programs". Socio-Econ. Plan. Sci., Vol. 18, No. 6.
- Campbell, D.T., (1957). factors Relevant to the Validity of Experiments in Social Settings. Psychological buletin., 54, 297-312
- Campbell, D.T., (1986). "Relabeling Internal and External Validity for Applied Social Scientist". New Direction for Program Evaluation., 31, 67-77.
- Campbell, D.T., and Stanley, J.C., (1966). Experimental and Quasi Experimental Design for Reseach. Skokie, Ill: Rand McNally
- Caws, P. (1965). The Philosophy of Science, Princeton, D. Van Nostrand Co.
- Copi, I.M., (1961). Introduction to Logic. The Mcmillan Company.
- Cook, T., and Campbell, D.T., (1976). "The design and conduct of quasi-experiments and true experiments in field settings". in: M. Dunnette (ed.) Handbook of Industrial and Organizational Psychology. Skokie, Ill.:Rand McNally,
- Cook, T., and Campbell, D.T., (1979). Quasi Experiments: Design and Analysis for Field Settings. Chicago: Rand McNally.
- Cordray, D.S., (1986). Quasi-Experimental analysis: A Mixture of Methods and Judgment. New Directions for Program Evaluation, 31 9-28.
- \*Cronbach, L.J. (1963). Cours improvement through evaluation, Teachers College Record, 64, 1963. Pp.
- Cronbach, L.J., (1983). Designing Evaluation of Educational and Social Programs. San Francisco: Jossey Bass

Dillon, G. (1981). Constructing Text: Elements of a Theory of Composition and Style. Bloomington, In: Indiana University Press.

\*Drewett, A. (1997), Evaluation and consultation, Learning the lessons of user involvement. Evaluation. 3(2) 189-204. Sage Publications, London.

Eyken, W.V., Goulden, D., Crossley, M., (1995). "Evaluating Educational Reform in a Small State: A Case Study of Belize, Central America." Evaluation, 1 (1), Sage Publication.

Fischer, W, Time and Chronic Illness. A Study on Social Constitution of Temporality. Habilitation Thesis, University of California, Berkeley.

Finne, H., Levin, M., Nilssen, T., (1995). "Trailing Research: A Model for Useful Program Evaluation." Evaluation, 1 (1), Sage Publication.

Glaser, B. & Strauss, A.L. (1967). The Discovery of Grounded Theory: Strategies for Qualitative Research. Chicago. Adline.

Geertz, C. (1973). The Interpretation Of Cultures, New-York, Basic Books,

Guba, E.G, and Y.S. Lincoln, (1989). Forth Generation Evaluation. London: Sage.

Hempel, C.G., (1965). Aspects of Scientific Explanation and Other Essays in the Philosophy of Science. New-York, The Free Press

House, E.R. (1980). Evaluating with validity. Beverly Hills, California, sage.

Hous, E.R., Editor's Notes in: Hous, E.R., (1983). Philosophy Of Education, New Direction For Evaluation. Jossey Bass

James, P.D., (1989) An Unsuitable Job for a Woman. Penguin Books, England.

Joint Commitee on Standards for Educational Evaluation, (1981). Standards for Evaluations of Educational Programs, Projects and Materials. New York: McGrew-Hill

Kazi, M.A.F., (1996). "Single-Case Evaluation in the Public Sector Using A Combination of Approaches." Evaluation. 2 (1), Sage Publication.

Krathwohl, D.R., (1982). Social and Behavioral Science Research: A Framework for Conceptualizing, Emplementing and Evaluating Research Studies. Jossey-Bass. San-Francisco

Krouglanski, A.M., Kroy,M. (1976). "Experimental Research: A Reconceptualization". Representative Research in Social Psychology ,7.

Kuhn,H. (1962). The Structure of Scientific Revolutions. Chicago: Universiyu of Chicago Press.

Lincoln, Y.S., Guba, E.G., (1986). "But is Rigorous? Trustworthiness and Autenticity in Naturalistic Evaluation. New Directions for Program Evaluation, june, 73-81.

Mark, M.M., (1986). "Validity Yypologies ans the Logic of Quasi Experimentation". New Directions Program Evaluation. Jossey Bass Inc. Publishers, 47-67.

Mor, Y. (1992). "The Use of Grey Documantation in The Comprehensive Evaluation of Broadly-Aimed Social Program". Paper presented at the International Expert Meeting on Information and Social Welfare Policy and Planning. November, Jerusalem.

Patton, M.Q., (1981). Creative Evaluation. Sage Publications,,: Beverly Hills.

Pawson R. (1996). Three steps to construcivist heaven. Evaluation. 2(2) 213-219. Sage Publication. London.

Peirce, C.S. (1960). Collected papers of Charles Sanders Pierce. (C. Hartshorne & Paul Weiss, eds.) Vol. III; IV, Cambridge, Belknap Press.

Popper, K.R., (1959). The Logic of Scientific Discovery. London, Hutchinson and Co. New-York, Basic Books, Inc.

- Provus, M. (1969). "Evaluation on Ongoing Programs In the Public School System" in: Educationa Evaluation: New Roles, New Means. R.W Tyler (ed.) Chicago; N.S.S.E.
- Provus, M. (1971). Discripancy Evaluation. Berkley, McCutchan Publishing Company
- Renzulli, J.S., (1975). A Guidbook for Evaluation Programs for the Gifted and Talanted. Office of the Ventura Count-Superintendent of Schools. Ventura, California.
- Rossi, P.H. and Freeman, H.E., (1982) .Evaluarion. Systemic Approach. Sage Publication, Beverly Hills, London, New Delhi.
- Rossi. P.H. and Right, J.D. (1984). "Evaluation Research, an Assesmen"t. Annual Review of Sociology. 10.
- Scriven, M., (1967). "The Methodology of Evaluation". in: R.E., Stake (ed.) Curriculum Evaluation, AERA Monograph Serues on Evaluation, No. 1 Chicago: Rand McNally.
- Scriven, M. (1972). "Pros and Cones about Goal-Free Evaluation".Evaluation Comment 3(4)
- Scriven, M., "The Evaluation Taboo". New Direction in Program Evaluation. September, (1983). 75-82
- Smith, M.L. (1986). "The Whole is Greater: Combining Qualitative and Quantative Approaches in Evaluation Studies". New Directions for Projects Evaluation, 30
- Stake, R.E. (1969) "Language, Rationality and Assessment".in: H.W Beaty (ed.) Improving of Educational Assessment. Association of Supervision and Curriculum Developoment, Washington.
- Stufflebeam, D.L. "The CIPP model for Program Evaluation" in Evaluation Models: Viewpoints on Educational and Human Services Evaluation. G.E. Madaus, M Scriven &D.L. Stufflbeam (eds.) Boston: Kluwer-Nijhoff.

Stufflebeam, D.L., W.J. Foley, W.J. Gephart, E.G.Guba, R.L. Hammond, H.O. Merriman, and M.M. Provus. ΦΔΚ National Study Committee on Evaluation. ( 4th ed. 1974) Educational Evaluation and Decision Making. Itaska, IL: F.E Peacock Publishers.

Tucker, A,S. and Dempsey, J,V. (1991). "A Semiotic Model for Program Evaluation" The American Journal of Semiotic, Vol. 8, No. 4 73-103

Turner, J,H,. (1986 ).The Structure of Sociological Theory. The Dorsey Press, (4th ed.)

Tyler,R.W. (1942). "General Statments on Evaluation" Journal of Educational Research..30

Tyler, R.W., (1950). Basic Principles of Corriculum and instruction. Chicago, University of Chicago Press.

Wallace, L.W., (1969). Sociological Theory. Heinmann Educational Books Ltd. London

Werner,W. (1978) Evaluation: Sense-making of school programmes. Monograph Series No. 1. Vancouver: Vancouver Center for the Study of Coriculum and Instruction, Faculty of Education, University of British Columbia.