**פרק (4) טרם נמצא הסבר להכרתיות (consciousness )**

ההכרתיות עדיין לא קבלה הסבר מכניסטי (בנוסח ההסברים הרווחים במדעי הטבע) המתאר בצורה טובה ומספקת את הקשר בין ההכרתיות לבין הנוירופיזיולוגיה של המוח, בדומה לקשר בין טמפרטורה לתנועה של החלקיקים הזעירים בחומר או בדומה לקשר בין אנרגיה לבין מאסה. בחלק הראשון של הפרק מועלות כמה השלכות מוזרות ובלתי הגיוניות העולות מההנחה שאכן נמצאה תיאוריה המתארת את הקשר בין הכרתיות לנוירופיזיולוגיה של המוח. בחלק השני מובאת סקירה קצרה של מספר ניסיונות להסביר את הקשר מוח-הכרתיות שעל בסיסן ניתן להגיע למסקנה שטרם נמצא הסבר לבעיית ההכרתיות.

הפרק מתחלק לשני חלקים. בחלק הראשון, אצא מההנחה שהמדע אמנם הצליח לפתח תיאוריה המסבירה את הקשר בין המוח להכרתיות ואראה שההשלכות מתיאוריה זו הן מוזרות מבחינה אינטואיטיבית ושתיאוריה זו אינה מתמודדת כהלכה עם מספר תצפיות אמפיריות. בחלק השני, אסקור בקצרה מספר ניסיונות לפתח תיאורטיות לבעיית המוח-הכרתיות ואראה שכולן אינן מניחות את הדעת.

1. **מהן ההשלכות העולות מההנחה שהקשר הכרתיות-נוירופיזיולוגיה אכן נתגלה?**

הבה נניח שנמצא הסבר להכרתיות של האדם, שהתגלה המכניזם הנוירופיזיולוגי היוצר או מאציל את החוויה הסובייקטיבית של ההכרתיות, כאשר היחיד רואה אדום או ירוק, טועם עוגה, וקורא עיתון, או שמכניזם נוירופיזיולוגי זה מתאפיין במיוחד בכך שאחת מתכונותיו הינה ההכרתיות. הסבר זה, שיקרא "חוק-ההכרתיות", הינו בעל צורה כללית של משוואה האומרת את הדבר הבא:

1. **Consciousness (C) = f(Brain's Neurophysiological Activity (BNA))**

מהן ההשלכות שאפשר להסיק מחוק זה, מהמשוואה: C = f(BNA)? אני אתרכז על ההשלכה הבאה, שאקרא לה "השלכת המדידה", המושתתת על השאלה הבאה: כיצד נמדוד את המושגים המופיעים בנוסחה זו? את ה BNA מודדים על ידי יחידות מכניסטיות סטנדרטיות, כמו למשל, על ידי הבדלים במתח החשמלי, עוצמת הזרם החשמלי, או מדידות של ריאקציות כימיות במוח. לשם קיצור, נקרא למדידות אלו, שכאמור מבוטאות ביחידות מדידה סטנדרטיות, בשם הכללי "יחידות סטנדרטיות" [Standard Units (su)]. ועתה, היות וההכרתיות שווה ל BNA, ומשום שבינתיים לא ידוע לנו על שום דרך פומבית למדוד הכרתיות כפי שמודדים מושגים שונים בפיזיקה או בכימיה (פרט כמובן לדיווח של כל אדם ואדם על מצבו הפנימי), נבטא את ההכרתיות על ידי מדידות סטנדרטיות, כלומר, בעזרת su המופיעים באגף ה BNA של חוק-ההכרתיות. גישת מדידה זו מבוססת על הדרישה ל "unit equivalency" (המושתתת על הניתוח הידוע של, dimensional analysis). לפי דרישה זו, הקומבינציה של יחידות המדידה בצד אחד של המשוואה (המבטאת חוק, תיאוריה) צריכה להיות זהה לקומבינציה של יחידות המדידה בצד השני של המשוואה (see Rakover, 2002, 2018). במילים אחרות, המסקנה הראשונה מחוק-ההכרתיות, C= f(BNA), היא שאת ההכרתיות ניתן למדוד על ידי יחידות מדידה סטנדרטיות, כלומר, ה su. (מן הראוי לשים לב שהיות וטרם נמצאה דרך פומבית למדוד הכרתיות באמצאות של יחידות מדידה מתאימות, יהיה קשה מאוד לבנות קבוע מסוים שיאזן את יחידות המדידה בשני האגפים של המשוואה: C = f(BPA) לפי הדרישה של unit equivalency.)

בהנחה שגישת מדידה זו היא אכן נכונה, ניתן להפיק מימנה מספר השלכות ומסקנות מאוד מעניינות אך מוזרות. כאן עלי להוסיף ולומר, שהשלכות מוזרות אלו הינן בחזקת ניסויי מחשבה, כלומר, אפשרויות העלולות להתרחש במציאות.

1. לפי המודל הנוכחי, אם ההכרתיות היא זו שמאצילה *משמעות* על ייצוגים מנטליים שונים, על ה MSs, המסמלים דברים שונים בעולם, רגשות ומחשבות שונות ומסובכות; ואם ההכרתיות היא זו שגם שמרכישה *הבנה* לכל ההסברים השונים, בין שהם מנוסחים באופן דדוקטיבי, אינדוקטיבי או בין שההבנה עולה כתוצאה מתפיסה של יחסים שונים בין מאורעות שונים, בכל המקרים האלה, יהיה אפשר לבטא את המשמעות ואת ההבנה על ידי ה su! למשל, אם אהבתו של אורי לאשתו היא בעלת משמעות גדולה, אזי נתן יהיה לבטאה ולציין שעבור אורי משמעות אהבתו לאשתו היא 10su; אבל אז עלול להתברר שהחשיבות שהוא מייחס לקניית מכונית משומשת היא גם בת 10su, ואילו המשמעות של המחשבה שהוא הולך לראות משחק כדורגל בין הפועל למכבי שווה 15su (זהירות! משמעויות שאשתו של אורי עשויה להיחשף להן בקלות רבה כי הן פומביות). כלומר, כאן עולה האפשרות הלא מקובלת שהמשמעות של דברים שונים לחלוטין, של איכויות מסוגים הזרים זה לזה, יקבלו במקרים רבים אותה חשיבות ושניתן יהיה לדרג אותם על אותו סולם המושתת על ה su. ובאותה צורה, ניתן יהיה לומר שיתכן מצב שבו מידת ההבנה של בעיה בתורת הקוונטים עשויה להיות זהה למידת ההבנה איך להכין חביתה עם בצל – מצבים די מוזרים, הייתי אומר.
2. נניח שבעיני הגברת סָלומון לפסל הבטון בעל הצורה של אישה שמנה ומכוערת מאוד הניצב בפתח ביתה יש משמעות גדולה השווה ל 23su. ועתה היות ויחידות אלו הן סטנדרטיות, ניתן לתרגם אותן בעזרת סידרה של נוסחאות טרנספורמציה ידועות היטב למדידות פיזיקליות מסוגים שונים. כתוצאה מתהליך זה, נמצא בסופו של דבר שבכפר מסוים בסין יש אריג ישן וצבעוני שניתן לייחס לו ערך של משמעות השווה ל 23su. וכאן עולים שוב מצבים מוזרים ביותר. ראשית, איך יתכן שאותה משמעות בת 23su שהגברת סלומון מייחסת לפסל הבטון שלה, אותה משמעות תייחס גברת סלומון גם לאריג הצבעוני בסין, כאשר הגברת סלומון עצמה כלל אינה מודעת לא רק לקיומו של האריג הצבעוני אלא גם לקיומו של כפר זה בסין. שנית, גם אם נטיס את הגברת סלומון ונראה לה את האריג הצבעוני, היא עלולה לעקם את האף ולומר, 'זה ממש מכוער!' ובכן, לא רק שלפי הנוסחה C = f(BPA) ניתן ליחס משמעות אישית של אדם ספציפי לכל מיני דברים וחפצים בעולם שאותו אדם ספציפי אף לא מודע לקיומם, אלא שגם הניבוי העולה מחוק-ההכרתיות הזה, מופרך כי אין התאמה בין הניבוי לתצפית! לפי החוק עולה הניבוי שבעיני הגברת סולומון משמעות האריג הצבעוני היא חיובית גבוהה, בעוד שהתצפית האמפירית מראה שמשמעות האריג עבור הגברת סלומון היא שלילית גבוהה.
3. אם ניתן למדוד הכרתיות, משמעות, והבנה על ידי שימוש ב su, אפשר יהיה לפתח גלולות המגבירות (מקטינות) את ההכרתיות, ואת המשמעות. למשל, דיקטטור מסוים עלול להכריח את נתיניו לקחת כל יום כדור אחד כדי להגביר את מידת משמעותו בעיניהם, וכדור שני כדי להקטין את מידת ההבנה של נתיניו ובכך להעצים את מידת הטמטום שלהם. זה מצב מוזר ללא כל ספק. יתר על, בעזרת הגלולות הללו ניתן יהיה לפתח מספר גאונים המתוכננים לממש את מטרותיו של השליט, ואוסף גדול של נתינים נרצעים כדי לבצע את כל העבודות המלוכלכות עבור משכורת מינימאלית מבזה.
4. בהתאם לשתי המסקנות (א) ו (ב) הנ"ל, אפשר לומר את הדברים הבאים: משבר משמעות-חיים אדיר שישנה את פני האנושות כולה, עלול להתרחש משום שעולם המשמעות וההבנה הפרטי והאישי של כל אדם ואדם יחדל מלהתקיים! ניתן יהיה לייחס להכרתיות (שמאצילה משמעות והבנה) מספר מסוים על הסקלה של su – אפשרות העומדת בניגוד גמור לכך שעולמו הפרטי של האדם פתוח רק לחוויה, התבוננות, האישית הבלעדית שלו. יתר על כן, בגלל שההכרתיות מועברת בעזרת חוק-ההכרתיות מהרמה הפרטית האישית לרמה הפומבית, אפשר יהיה לטעון שגם רובוטים יוכלו לחוות חוויות אישיות בדומה לגברת סלומון. למשל, הרובוט R300 יחוש משמעות בת 23su לפסל הבטון ועלול אף לעקם את האף למראה של האריג הצבעוני בכפר הסיני. ולא רק זאת, אם התפיסה עד לתגלית חוק-ההכרתיות הייתה שמשמעות מיוחסת לעולם אדיש על ידי האדם, אפשר יהיה להעלות את האפשרות שלכל החלקים בעולם יש ערכים משמעותיים המיוחסים להם בעזרת מערכת חוקי הטרנספורמציה הידועים למדע. האם נובע מכך, שהמשמעויות של חלקי התבל הן אובייקטיביות ובלתי תלויות בהערכה האנושית?

היות ותשובה חיובית מובילה למסקנה על אודות עולם מופלא בעל משמעות-חיים כתכונה טבעית ואובייקטיבית, והיות שזו מסקנה הנראית בעיני כל כך מהפכנית ועד כדי כך משונה שאני מעדיף לעבור מיד לדון בתשובה השלילית. בהנחה של תשובה שלילית ובהנחה שעדיין נקבל את הטענה שהאדם הוא זה שמייחס משמעות לעולם אדיש, תיווצר כאן בעיה מסובכת מאוד. מדוע? משום שהמשמעות של כל חלק בתבל עשויה להשתנות בהתאם למשמעות האינדיבידואלית של כל אחד מבני האדם המאכלסים את כדור הארץ. למשל, המשמעות של השמש עשויה להשתנות מעבר למיליארדים של בני תמותה. לא ברור איזה ערך של משמעות ידבק בשמש, זה של X או זה של Y או שמא ניתן יהיה לחשב ממוצע של משמעות השמש מעבר לכל בני האדם. מכל מקום, יהיה עלינו להתחשב בכך שגם לבעלי חיים (חתולים, כלבים, קופים, סוסים, חמורים, כבשים ופרות) יש רמה מסוימת של הכרתיות ועל מספר חלקים מהעולם מואצלת איפה משמעות חיתית.

אני סבור שמה שנאמר לעיל דיו כדי לעורר חשד שחוק-ההכרתיות מעורר שלל של בעיות המטילות ספק בתכנית המחקר שמטרתה לגלות תיאורית מוח/הכרתיות בעלת מבנה מהסוג של C = f(BPA). איך ניתן להגיב למסקנות המוזרות העולות מחוק זה? הינה מספר אפשרויות להתמודדות:

1. חוקרים ייטו לחפש ליקויים בלוגיקה של המסקנות המוזרות המובאות לעיל. אם ימצאו ליקויים כאלו, הרי שהמטרה לפתח תיאוריה בנוסח של חוק-הכרתיות תקבל חיזוק.
2. אם לא ימצאו ליקויים שיחסלו את המסקנות המוזרות, חוקרים עשויים לומר את הדברים הבאים. ראשית, המסקנות הן במהותן אמפיריות, כלומר, אלו מסקנות הקשורות לעולם התצפיתי שמטבעו הוא בלתי צפוי ומכיל אינסוף הפתעות חשובות ומעניינות. לכן, ניתן לומר שהמעמד של המסקנות האלו אינו דומה לסוג ההוכחות המתמטיות או ההנדסיות, כמו למשל, הטענה שבהנדסת המישור אין משולש שסכום הזוויות שלו קטן מ 180 מעלות. אם יימצא חוקר שיעשה מאמץ עליון כדי להוכיח שבהנדסה של אוקלידס יש משולש מוזר כזה, יעורר הדבר תמיהה גדולה מאוד (כמובן, שישנה הנדסה לא אוקלידית שבה יש משולש כזה). אולם, כאמור, זה לא המצב לגבי חוק-ההכרתיות! לכן, אפשר להציע את האסטרטגיה הבאה: הבה נעשה מאמץ עליון, נגלה את המכניזם המקשר בין הנוירופיזיולוגיה של המוח לבין ההכרתיות, ולאחר מכן נדאג לכל אחת מהמסקנות השליליות הנ"ל ובוודאי שגם לעוד כמה מסקנות שליליות העלולות להתגלות במשך הזמן.

שנית, חוקרים עשויים להציע שאולי תכנית המחקר של העמדת ההכרתיות על הנוירופיזיולוגיה של המוח הגיעה למבוי סתום, ואולי הגיע הזמן לחפש דרכים חלופיות להסברת ההכרתיות – דרך שאני לא מצאתיה בספרות המקצועית. (אולי כדאי להתייחס אל ההכרתיות, כפי שאציע בסוף הפרק, כאל גורם מסביר בסיסי וראשוני.)

**(2) סקירה קצרה על הניסיונות לפתח תשובות לבעיית המוח וההכרתיות**

סקירה זו תוביל אותי להתרשמות הבאה: עד כה טרם הצליחו לפתח תיאוריה המסבירה את ההכרתיות בעזרת תהליכים נוירופיזיולוגיים במוח, "תיאורית הכרתיות-מוח", בדומה לתיאוריות במדעים המסבירות טרנספורמציות של אנרגיה, כמו חיכוך וחום, אנרגיה פוטנציאלית וקינטית, חשמל ומגנטיות, אנרגיה ומאסה; או תיאוריות המסבירות כיצד חומרים משתנים כמו מים הנוצרים ממזוג של גז החמצן עם גז המימן וכיצד אלקטרוליזה מפרקת מים לשני גזים אלה.

כדי לבסס את התרשמותי זו, אפרק את השאלה הכללית: האם ניתן לתפוס את החוויה-הכרתית במושגי מדעי הטבע, לארבע תת-השאלות הבאות:

1. האם ניתן להסביר התנהגות אנושית באופן מכניסטי בלבד מבלי להשתמש במושג ההכרתיות כמושג מסביר?
2. האם הצליחו לפתח תיאוריה המסבירה את הקשר בין תהליכים נוירופיזיולוגיים במוח לבין הכרתיות?
3. האם ניתן להעמיד (reduce) הסברים מנטליסטיים המעוגנים במושג ההכרתיות על הסברים מכניסטיים (כמו למשל, על הסבר נוירופיזיולוגי)?
4. האם הצליחו לפתח מחשב מאוד מסובך ומתוחכם המסוגל כמונו להימצא במצב של הכרתיות?

תשובה חיובית על אחת מארבע שאלות אלו פירושה שאין צורך בהכרתיות כדי לטפל בהתנהגות אנושית (והתנהגות של בעלי חיים עילאיים), משום שבמקרה זה תתאים המתודולוגיה הסטנדרטית שפותחה במדעים למחקר ולהסבר של כל סוגי ההתנהגויות שבהן דנה הפסיכולוגיה. רק תשובה שלילית לכל השאלות האלו מחייבת דיון רציני וגם פיתוח גישה הסברית חדשה המתאימה לתחומי המחקר הפסיכולוגיים הקשורים בהכרתיות, הבנה, ומשמעות.

1. *האם ניתן להסביר התנהגות אנושית באופן מכניסטי בלבד מבלי להשתמש במושג ההכרתיות?*

חוקרים רבים האמונים על הגישות של הבהביוריזם, הפסיכולוגיה קוגניטיבית, והפסיכולוגיה הקוגניטיבית-נוירופיזיולוגית, סבורים שכל התנהגות ניתנת להסבר בהתאם למודלים מכניסטיים של הסבר, מודלים המקובלים במדעי הטבע, ושאין שום צורך להשתמש במושג ההכרתיות כגורם המסביר את ההתנהגות (ראו, (Rakover, 2018. אקרא לגישה זו "אי-נחיצותה של ההכרתיות" (consciousness despensibiliuty). ההתרשמות שלי היא שבבסיס גישה זו עומד הטיעון הבא, שאקרא לו "ריבוי-הסברים": לכל קבוצה של תוצאות המוצגת כקבוצה של נקודות במערכת צירים של Y=f(x), אפשר להתאים מספר אינסופי של פונקציות מתמטיות שינבאו היטב קבוצת נקודות זו, כשכל פונקציה מהווה הסבר מכניסטי להתרחשותן של התוצאות בתנאים מסוימים. בהינתן טיעון זה, יוצא שבאופן עקרוני אין תופעה פסיכולוגית שלא ניתן להציע לה הסבר בנוסח של מדעי הטבע, ושלאמתו של דבר עולה מכאן המסקנה שהשימוש במושג ההכרתיות כמושג מסביר מיותר לחלוטין. להלן מספר ציטוטים התומכים בטיעון זה של "אי-נחיצותה של ההכרתיות":

למשל, זוג החוקרים Radner & Radner (1989) כתבו:

"When animal consciousness is dismissed as superfluous, we must ask whether the dismissal refers to consciousness as a phenomenon to be explained or as an explanatory device. The most plausible answer is that consciousness is superfluous in the latter role. Anything that can be explained by it can be explained equally well without it." (p. 206).

ועל מושג ה'מטרה' המנטליסטי כתב Keijzer (2001) את השורות הבאות:

"The goal is to formulate an explanation which does not involve any thinking or sentient agent in its premises. The explanans should involve no one who is acting as an intelligent, sentient force, guiding behavior in the right direction." (p. 26)

דברים דומים כתב Dawkins (1995)::

"There is no prediction we can make that if the animal has consciousness it should do X but not conscious it should do Y" (p. 139).

וכך גם Flanagan (1992)שהציע את המושג "conscious inessentialism" כדי לאפיין את אי-נחיצותו של מושג ההכרתיות:

"… the view that for any activity i performed in any cognitive domain d, even if we do i consciously, i can in principle be done nonconsciously" (p. 129).

לאמתו של דבר ניתן להסיק משתי הגישות הפילוסופיות: ה *identity theory*, שמציעה שקיימת זהות בין MS לבין neurophysiological state (NS), וה Functionalism שמציעה ש MS מוגדר בצורה פונקציונאלית, באופן תפקודי, וניתן למימוש בצורות שונות (ראו למשל Robb & Heil, 2019), מסקנה זהה: את התפקיד ההסברי לא נושא המושג המנטליסטי, ה MS, אלא המושג הנוירופיזיולוגי, ה NS, שאותו קל לקשר באופן סיבתי אל ההתנהגות האדם או בעל החיים. במילים אחרות, ההסבר להתנהגות ניתן בעזרת מודל סיבתי מקובל שבו המושג המסביר את ההתנהגות הוא ה NS ולא ה MS, פשוט משום שעדיין לא ברור כיצד ה MS מתקשר למאורעות נוירופיזיולוגיים שונים במוח (כלומר, לבעיית הגוף/נפש טרם נמצא פתרון).

אני דוחה לחלוטין את גישת אי-נחיצותה של ההכרתיות מהסיבות הבאות:

ראשית, אני דוחה גישה זו משום שללא מושג ההכרתיות נתקשה מאוד להבין הן התנהגות אנושית והן התנהגות של בעלי חיים, משום שהתנהגויות אלו רוויות בהכרתיות, כמו למשל, רצון, אמונה, שאיפה, כוונה וכו ולא ניתן להתעלם מכך. בספרי Rakover (2007)*, To understand a cat: Methodology and philosophy*, אני מתאר מספר רב של אפיזודות התנהגותיות הקשורות ביחסים בין חתול הבית, מקס, לבין אביבה אשתי או לביני, שקשה מאוד להסבירן מבלי לשייך לחתול רצון, מטרה, וכוונה. כלומר, קשה מאוד להבין את האינטראקציה בין מקס לביננו רק בעזרת הסברים ביהביוריסטיים, אינסטינקטיביים, או למידתיים אוטומטיים – הסברים מכניסטיים העולים בקנה אחד עם המתודולוגיה הרווחת במדעים.

שנית, אני דוחה את גישת 'אי-נחיצותה של ההכרתיות', משום שהטיעון הנ"ל, 'ריבוי-הסברים', מבוסס על הנחה סמויה ומוטעית. הנחה זו קשורה במתודולוגיה התצפיתית של הפסיכולוגיה. מה שהפסיכולוגיה עושה כשהיא מספקת הסברים לתצפיות על התנהגויות בתנאים מסוימים, הוא להציע הסברים לתצפיות פומביות, כלומר, לתצפיות שכל אחד יכול לעשותן לפי הדרישה המתודולוגית הנקוטה במדעים ושאותה אימצה הפסיכולוגיה (ראו (Rakover, 1990. תצפית התנהגותית זו אינה אלא התנהגות שהופשטה ממנה כל משמעות אישית, כל אלמנט של הכרתיות. האינדקסים של ההתנהגות, כלומר, המספרים שהפסיכולוגים מייחסים להתנהגויות הנצפות, שעליהם עורכים הפסיכולוגים ניתוח סטטיסטי ושאותם הם מדווחים בירחונים המקצועיים, אינם מייצגים את המטרה, המשמעות והכוונה האישית של האדם, אלא את התגובה ברת התצפית הפומבית (למשל, תנועה מוטורית) שביצע המשתתף בניסוי. למשל, האינדקסים של 'אחוז נכונים' (Percent correct) ו'זמן חביון' (Reaction time) הם מדדים שאינם לוקחים בחשבון את המטען ההכרתי של המשתתף בניסוי, אלא רק את ביצוע התגובה המוטורית שלו, כלומר, אם הוא/היא לוחץ על הכפתור של התגובה הנכונה או לא, וכמה זמן חלף מזמן הצגת הגירוי עד למתן התגובה. אלה הם תגובות שרובוט מתוחכם מסוגל לבצע גם מבלי שיהיה ניחן ביכולת להכרתיות (הרי הוא בסך הכול מכונה). למשל, תוכנה משוכללת יכולה להחליט אם תמונת הפנים המוצגת לפניה מצויה או לא מצויה במאגר תמונות הפושעים. באותו אופן, עד אנושי יכול לעבור על תיקים במשטרה ולזהות את פני הפושע שהוא ראה לא מכבר באירוע של פשע. בשני המקרים נקבל אותה תוצאה: או שהזיהוי הצליח או שלא. אולם, וזו השאלה הקריטית, האם התוכנה לזיהוי פלילי מבינה כבן אנוש את מעשיה ואת השלכותיה? איני סבור שכן. בהנחה שאכן ניתוח זה נכון, אפשר לומר את הדבר הבא: משום שהאינדקסים התנהגותיים אינם רווים בהכרתיות, לא עולה הדרישה להסבר המושתת על מושג ההכרתיות. אי לכך, הסברים מכניסטיים (המשוללים כל נגיעה להכרתיות) עשויים להציע הסברים מספקים למדי לאינדקסים מסוג זה, אינדקסים אובייקטיביים. אולם, יש לציין, שהסבר מסוג זה ניתן להתנהגות משוללת כל משמעות הכרתית – הוא יתאים באותה מידה גם להתנהגותו של רובוט או זומבי.

ושלישית, הספרות הפילוסופית רצופה הצעות לפתרון בעיית הגוף-נפש, המוח-הכרתיות, שכולן בלי יוצא מהכלל קיבלו ביקורת חריפה ביותר (ראו למשל, Kim, 1996; Rakover, 2018; Robb & Heil, 2019). אני אתן כמה דוגמאות להמחשה. דוגמה ראשונה, ב 1982 פרסם Jackson מאמר מפורסם המראה שגם אם נדע את כל שיש לדעת, למשל, על ראיית צבעים, לא נוכל להסביר בעזרת הידע הזה את חווית ראית הצבע האדום. ג'קסון סיפר סיפור מעניין על מדענית מומחית לראיה, שידעה כל שניתן לדעת על הפיזיקה, הכימיה, הפיזיולוגיה והפסיכולוגיה של ראיית צבעים. למדענית הזו קראו מרי והיא חיה כל ימי חייה בסביבה של שחור-לבן. יום אחד יצאה מרי מסביבתה השחורה-לבנה וראתה, חוותה, בפעם הראשונה בחייה צבע אדום. היא למדה משהו חדש שלא סופק לה קודם על ידי הידע המדעי המושלם שלה, היא למדה מה פירוש הדבר לראות צבע אדום, לחוות את התחושה ההכרתית של ראיית צבע אדום.

דוגמה שנייה, ב 1977 פרסמו Nisbett & Wilson מאמר מפורסם הטוען שבניסויים רבים בפסיכולוגיה חברתית ובעשיית החלטות התברר שהמשתתף בניסוי לא היה מודע לגירוי, לתגובה, להשפעת הגירוי על התגובה או לתהליכים עליונים רלוונטיים. הטיעון התבסס על הממצא שלפיו ההסברים שנתנו המשתתפים בניסויים להתנהגותם הם היו מוטעים. המאמר עורר ביקורות רבות, וכדוגמה אחת אתאר בקיצור נמרץ את הביקורת שלי (ראו Rakover, 1983). המשתתפים בניסויים של Nisbett & Wilson היו אכן מודעים לתהליכים שהתרחשו בראשם, אולם כמו כל אנשי המדע, ברוב המקרים הם הציעו את ההשערה המסבירה הלא נכונה, כלומר, הם בחרו מכל ההסברים שהתרוצצו בראשם את ההסבר המוטעה. לכן, לא עולה בהכרח מכך שהמשתתף לא היה מודע.

ודוגמה אחרונה, Velmans (1991) הציע שבמקרים רבים כמו בדיבור, היחיד מודע להתנהגות המקרים לאחר שהיא בוצעה. על סמך זאת הוא הסיק שלהכרתיות חשיבות מועטה ביותר בתהליך של עיבוד האינפורמציה ושתהליכים לא-הכרתיים הם אלה המתווכים בין הופעת הגירוי לבין תגובתו של המשתתף בניסוי. מאמר זה עורר ביקורת ודיון נרחב בנושא. לדעתי הטיעון של ולמאנס אינו נכון, כי להכרתיות חשיבות עצומה בעיבוד האינפורמציה המתווכת בין הגירוי לתגובה (ראו Rakover, 1996). הצעתי ניסוי מחשבתי ( The mental-pool thought experiment) שהתבסס על הידע הפסיכולוגי הקיים שלפיו כמות האינפורמציה ההכרתית הוא מוגבלת בעוד שכמות האינפורמציה הלא-הכרתית אינה מצומצמת. הצעתי שהאינפורמציה מהגירוי מעובדת תחילה ברמה הלא-הכרתית אולם לאחר מכן היא עוברת לרמה ההכרתיות שמאפשרת את הופעתה של התגובה ואת הפיקוח עליה.

1. *האם הצליחו לפתח תיאוריה המסבירה את הקשר בין תהליכים נוירופיזיולוגיים במוח לבין ההכרתיות (תיאורית מוח-הכרתיות)?*

על בסיס של סקירה רחבה ומעמיקה של הניסיונות לבנות תיאורית מוח-הכרתיות, הגיעו Cosmelli, Lachaux & Thompson (2007) למסקנה שכל מה שהצליחו לעשות עד כה הוא למצוא לא הסבר נוירופיזיולוגי להכרתיות, אלא רק אסוציאציות, קורלציות בין מדדים נוירופיזיולוגיים לבין מדדים של הכרתיות:

"…the neurodynamical approach works at the level of correlations, albeit refined ones." (p. 763)

חוקר ההכרתיות הידוע Chalmers (1996, 1997a) הציע להבחין בין שני סוגים של בעיות בהכרתיות: הקלה והקשה, וכך הוא כותב:

"The easy problems of consciousness are those that seem directly susceptible to the standard methods of cognitive science, whereby a phenomenon is explained in terms of computational or neural mechanisms. The hard problems are those that seem to resist those methods." (1997 p. 9)

לדעתו, הבעיות הקלות קשורות בהסברים של התנהגויות, כגון, תגובה לגירוי, עשיית הבחנות, מיקוד תשומת הלב, ארגון של אינפורמציה, דיווח ורבלי של מצבים מנטאליים, ושליטה רצונית על התנהגות. התנהגויות מסוג זה ניתנות להסבר על ידי פנייה למכניזמים קוגניטיביים ונוירופיזיולוגיים, כלומר, בעזרת מכניזמים המממשים את הפונקציות של התנהגויות אלו, כמו למשל, מציאת המימוש לעשיית הבחנה בין גירוי A לבין גירוי B. וכך הוא כותב:

"To explain access and reportability, for example, we need only specify the mechanism by which information about internal states is retrieved and made available for verbal report." (1997 p. 10)

הגישה של צ'למארס עוררה התנגדות רבה אצל מספר חוקרים (ראו אצל Shear, 1997). גם אני איני מקבל גישה זו, משום שלדעתי רוב התנהגויות רוויות בהכרתיות ולכן מתן הסבר לבעיה הקלה הוא לא פחות קשה מאשר מתן הסבר לבעיה הקשה. כפי שטענתי לעיל, ההסברים שהפסיכולוגיה הקוגניטיבית מספקת לבעיות הקלות הינם הסברים מכניסטיים המתאימים להתנהגויות שההכרתיות הוסרה מעליהם. במילים אחרות, ההסברים הניתנים לבעיות הקלות הינם הסברים הניתנים להתנהגויות של זומבים או של רובוטים נטולי הכרתיות, ולא להתנהגויות אנושיות שהינן ספוגות בהכרתיות.

כדוגמה אחרונה בסקירה קצרה זו של הסעיף הנוכחי, אדון בקיצור בתיאוריה של Giulio Tononi, ה Integrated Information Theory (IIT) of consciousness, שזכתה בשנים האחרונות, מצד אחד, להתעניינות רבה, ומהצד האחר, לביקורות חריפות (ראו למשל, Fallon, 2019; Tononi, 2015; Tononi, Boly, Massimini & Koch, 2016). כפי שנראה להלן, זו היא תיאוריה מהסוג של תיאורית-זהות (Identity Theory) והיא מבוססת על שלושת הרעיונות הבאים: ראשית, פירוט של תכונות ההכרתיות (הנתפסות כאקסיומות); שנית, פירוט של התכונות המצע הפיזיקאלי (הנוירופיזיולוגי) שנדרשות כדי לממש את התכונות ההכרתיות (הנתפסות כפוסטולטים); ושלישית, קביעה של זהות בין הכרתיות לבין סוג מסוים של אינפורמציה המעובדת על ידי המצע הפיזיקאלי ומניסוח פורמאלי של מדד המבטא את דרגת ההכרתיות (Φ).

ה IIT יוצאת מהעמדה שלפיה ההכרתיות היא תופעה פנומנולוגית קיימת; שלהכרתיות יש מבנה והיא מורכבת; שהכרתיות מכוונת לדברים מסוימים, שהיא נושאת אינפורמציה; שהכרתיות היא אחידה (אי אפשר לחוות את הצבע האדום של העגבנייה בנפרד מצורתה); ושלהכרתיות יש גבולות משום שהיא מכוונת לדבר מסוים ולא לאחר – תכונות שכל אחד מסכים עמן ולכן היחס אליהן הוא כאל אקסיומות של ההכרתיות. בהתבסס על תכונות אלו התיאוריה ממשיכה ומפרטת את התכונות שסיסטמה פיזיקאלית חייבת לגלות כדי שתוכל לממש את ההכרתיות (הפוסטולטים). למשל, על בסיס התכונה שלהכרתיות יש תכונה של נושאת אינפורמציה, עולה הפוסטולט שלפיו הסיסטמה הפיזיקאלית, הנוירופיזיולוגית, חייבת להיות מושתתת על אלמנטים המסוגלים ליצור קומבינציות, שילובים, מבנה הניחן בכוח של סיבה-תוצאה (combinations that have cause-effect power) המכוון למימוש מצב הכרתי ספציפי, לדבר מסוים.

כשאדם רואה חתול על הספה, לפי ה IIT, נוצר במוחו "מבנה-מושגי" (Conceptual structure) המבוסס על מספר מושגים מסוימים והיחסים בינם לבין עצמם, כלומר, אינפורמציה משולבת (integrated), מבנה המייצג את הנראה (החתול על הספה) והמטופל לפרטי הפרטים על ידי המערכת הפיזיקאלית, כלומר, "המצע-פיזיקאלי" (Physical substrate) שמתפקד בהתאם לפוסטולטים הנ"ל (כאמור, אצל האדם מצע זה הינו הנוירופיזיולוגיה של המוח). ה IIT מבוססת על הזהות היסודית בין החוויה ההכרתית (הראייה של החתול על הספה) לבין המבנה-מושגי של חוויה זו, כשמבנה-מושגי זה ממומש לפרטי הפרטים על ידי הנוירופיזיולוגיה של המוח. באופן כללי, אם כן, ה IIT מציעה שההכרתיות זהה לסוג מסוים של אינפורמציה משולבת, הממומשת על ידי מערכת פיזיקאלית ספציפית. מערכת זו ניתנת לחלוקה לתת-קבוצות שונות בעלות מבנים שונים של סיבה-תוצאה. תת-הקבוצה בעלת המבנה המקסימאלי של סיבה-תוצאה שאינו ניתן לרדוקציה לפרטים המרכיבים אותו, היא תת-הקבוצה המבטאת את המבנה המושגי המקסימאלי שלא ניתן לרדוקציה (Maximally irreducible conceptual structure (MICS)). לפי ה IIT מבנה זה, ה,MICS הינו המצב ההכרתי, ולכן ניתן לומר שהמערכת הפיזיקאלית שמקיימת את הפוסטולטים הנ"ל ניחנה באופן פנימי בתכונה של ההכרתיות כשם שמאסה ניחנה בתכונה של כוח המשיכה. מערכת זו היא מכניזם הפועל בצורה משולבת של סיבה-תוצאה ובכך היא מארגנת אינפורמציה. את דרגת המורכבות של ה MICS ניתן לייצג בעזרת מדד הנקרא Phi (Φ). ככל שה MICS מורכב יותר, כך גדלה דרגת ההכרתיות, כלומר, כך גדל ה Φ שמבטא באופן מספרי את ה MICS. כאשר MICS הוא מקסימאלי, מבטא גודלו של Φ את דרגתה המקסימאלית של ההכרתיות (במקרה זה הסימון הוא ΦMax).

ה IIT קיבלה חיזוק מכמה ממצאים אמפיריים. למשל, התברר שבשינה עמוקה המשוללת חלומות, אינדקסים הקרובים ל Φ מראים ערכים קטנים מאלה בשעת ערות. כמו כן נמצא בהתאם ל IIT, שפגיעה במוח הקטן (cerebellum) אינה פוגעת בהכרתיות משום שהתאים במוח הקטן אינם נמצאים באינטראקציה בינם לבין עצמם כפי שדורשים הפוסטולטים.

כאמור, ה IIT קבלה ביקורת רבה (ראו סיכום ודיון אצל Fallon, 2019). כאן אני חפץ להדגיש את נקודה הביקורת הבאה המתקשרת לאשר אמרתי לעיל: ברגע שמשתיתים בצורה זו או אחרת את ההכרתיות על הנוירופיזיולוגיה של המוח, באותו הרגע הופכת ההכרתיות לתכונה ברת מדידה פומבית בעזרת מדדים סטנדרטיים המקובלים במדעים וכתוצאה מכך עולים מסקנות מוזרות ביותר שרמזתי עליהם קודם לכן.

ראשית, עולה האפשרות לבנות מערכת מכנית שתעמוד בכל הדרישות של ה IIT, ולכן אפשר יהיה להציע שמערכת מכנית זו היא בעלת הכרתיות ועלול אף להיות מצב שבו ה Φ של מערכת זו יבטא הכרתיות גדולה מזו של האדם. אפשרות זו נוגדת לחלוטין את האינטואיציה ואת השכל הישר, משום שכל אחד יודע שמכונה היא מכונה ועד היום לא נתגלה בשום מכשיר כולל מחשבים מאוד מתוחכמים שמץ של הכרתיות. (זו ביקורת שנדונה גם אצל Fallon, 2019). התגובה של Tononi, Boly, Massimini & Koch, 2016 היא כפי שניתן לראות מעניינת במיוחד (הם מוכנים לקבל את האפשרות הזו), וכך הם כותבים:

"Intriguingly, IIT allows for certain simple systems such as grid-like architectures, similar to topographically organized areas in the human posterior cortex, to be highly conscious even when not engaging in any intelligent behaviour." (p. 460. See also p. 458).

שנית, אפשר להעלות טיעון נגד דרך השימוש במושג האינפורמציה על ידי ה IIT. היות שמושג האינפורמציה תלוי בהכרתיות של האדם, יוצא שהניסיון של ה IIT להבין הכרתיות בעזרת אינפורמציה הוא מעגלי (כי כאמור אינפורמציה תלויה בהכרתיות). התשובה לביקורת זו מבוססת על כך שמושג האינפורמציה לפי ה IIT מובנה לתוך המצע הנוירופיזיולוגי שמטפל במבנה-מושגי, וכך כותבים Tononi, Boly, Massimini & Koch (2016) :

"In IIT, information is causal and intrinsic: it is assessed from the intrinsic perspective of a system based on how its mechanism and present state affect the probability of its own past and future states (cause-effect power)." (p.457).

נדמה לי שתשובה מסוג זה המבוססת על כך שאינפורמציה מוערכת מפרספקטיבה מהותית של המערכת הנוירופיזיולוגית עצמה, מזמינה פירוש של איש קטן השוכן במערכת זו (Homunculus) ולכן בעיית ההכרתיות נשארת בעינה.

לסכום עניין זה של ה IIT, אם בביקורת הנוכחית ובאלו הנוספות שמופיעות בספרות המקצועית, יש אמת, הרי עולה מכך שגם התיאוריה הזו לא פענחה את חידת ההכרתיות. מה שהתיאוריה כן הצליחה לעשות הוא לאתר סוג מסוים של מבנה נוירופיזיולוגי במוח, שלשינויים במידת האקטיבציה שלו יש קורלציה גבוהה עם שינויים בדרגת ההכרתיות – וזה לדעתי מעניין מאוד.

1. *האם ניתן להעמיד (reduce) הסברים מנטליסטיים המעוגנים במושג ההכרתיות על הסברים מכניסטיים (למשל, הסבר נוירופיזיולוגי)?*

התשובה לשאלה זו היא מסובכת, כי קשה לערוך הקבלה בין רדוקציה תיאורטית (בין תיאוריה אחת לשנייה), נושא שנידון לעומק בספרות המקצועית, לבין רדוקציה בין מודלים של הסבר. מדוע? כי מודלים של הסבר אינם תיאוריות מדעיות בדומה לתיאוריות של תפיסה, למידה, וזכירה, אלא בעיקרם הינם פרוצדורות המנחות את החוקר כיצד להציע הסברים בתחומי מחקר מסוימים. לכן קשה מאוד לראות מהי הדרך הרציונאלית שבה ניתן יהיה להעמיד הוראות ביצוע מסוג אחד על הוראות ביצוע מסוג שני. היוצא מכך הוא שהתשובה לשאלה הנוכחית לא תתמקד במיוחד בהעמדת מודל הסבר מנטליסטי (כמו הסבר מטרתי, טלאולוגי) על מודל של הסבר מכניסטי (כמו הסבר סיבתי), אלא התשובה תתמקד יותר באפשרות של העמדה של תיאוריה מנטליסטית (המבוססת על מושגים סובייקטיביים הכרתיים שקשורים לעולמו הפנימי של היחיד, רצון, אמונה, כוונה, תחושות ורגשות, וכדומה) על תיאוריה מכניסטית (המעוגנת במושגים אובייקטיביים הקשורים בפיזיקה, כימייה, פיזיולוגיה, תהליכי מחשב וכדומה). הסיבה העומדת מאחורי מאמץ רדוקציוני זה נעוצה, בין היתר, בכך שמבחינה מתודולוגית הפסיכולוגיה מעדיפה הסברים מכניסטיים על פני הסברים מנטליסטיים, משום שהפסיכולוגיה השתיתה את עצמה על המתודולוגיה המדעית המקובלת במדעי הטבע. תחילה אעמוד בקיצור על הבעיות הקשורות ב psycho-neural reduction ולאחר מכן אטפל במאמצים אחרים לתפוס מושגים מנטליסטיים באמצעות מושגים מכניסטיים, כמו למשל, החלפת הסבר מנטליסטי (כמו הסבר מטרתי, טלאולוגי) בהסבר מכניסטי (כמו הסבר סיבתי המבוסס על תהליכים נוירופיזיולוגיים). הדיון בכל המקרים האלה יראה שעדיין לא נמצאה דרך ברורה ומקובלת להעמיד או להמיר הסבר מנטליסטי במכניסטי.

השאלה שאתחיל בה היא זו: האם אפשר להעמיד ((reduce תיאוריה פסיכולוגית, המבוססת על מושגים הקשורים בעולמו הפנימי של היחיד (רצון, אמונה, כוונה, מטרה, רגש), על תיאוריה נוירופיזיולוגית? זו היא השאלה המוכרת בספרות בשם של psycho-neural reduction (ראו למשל, Kim, 1998; Rakover, 1990; van Riel & Van Gulick, 2016). תחילה אסביר בקיצור מהי הפרוצדורה שבעזרתה מנסים להעמיד תיאוריה A על תיאוריה B. כדי לא להכביר מילים בנושא מסובך זה, אתאר רק את הגישה הקלאסית של :Nagel (1961) תיאוריה A, המכונה the reduced theory (TR), מעומדת על תיאוריה B, המכונה the basic reducing theory (TB), כאשר ניתן לגזור, להסיק, את TR מ TB יחד עם חוקי גישור המקשרים בין מושגי שתי תיאוריות אלו. חוקי הגישור נתפסים בדרך כלל כזהויות בין מושגי TB לבין מושגי TR. למשל, במקרה של העמדת התרמודינמיקה על הסטטיסטיקה המכנית, הוצע שמושג הטמפרטורה זהה למושג של האנרגיה הקינטית.

נגד האפשרות לערוך רדוקציה פסיכו-נויראלית, בין TR פסיכולוגית לבין TB נוירופיזיולוגית, הועלו מספר טיעונים. כאן אעמוד בקיצור על הטיעון המפורסם הנקרא ריבוי-מימושים (multiple realizations) (למשל, Fodor, 1974, 1998). נסתכל כדוגמה על המצב המנטלי 'כאב' [Mental Statepain (MSpain)]. מצב מנטלי זה, המוגדר מבחינה תפקודית (כי ל MSpain יש התנהגות ספציפית ופונקציה תפקודית ברורה, להפסיק את הנזק ליחיד), יכול להתממש בדרכים שונות ועל ידי תהליכים מוחיים נוירופיזיולוגיים שונים המופיעים אצל בני אדם, קופים, כלבים, חתולים, דגים וכו. יתר על כן, ה MSpain המייצג התנהגות של כאב ברובוט יכול להתממש על ידי חומרים שונים מזה של האדם. אם טיעון זה תופס, הרי יוצא שלא יהיה ניתן למצוא חוק גישור בין מושגי TB לבין מושגי TR, משום שניתן לממש את מושגי TR בדרכים שונות ובעזרת תהליכים שונים. ומכן עולה, שאם לא ניתן למצוא חוק גישור פסיכו-פיזיולוגי לא ניתן יהיה להגשים את השאיפה להעמיד תיאוריה פסיכולוגית על תיאוריה נוירופיזיולוגית.

טיעון נוסף נגד האפשרות לערוך רדוקציה פסיכו-נויראלית מעוגן בדרישה ל "unit equivalency" (שכאמור מושתתת על הניתוח הידוע של, dimensional analysis). לפי דרישה זו, הקומבינציה של יחידות המדידה בצד אחד של המשוואה (המבטאת חוק, תיאוריה) צריכה להיות זהה לקומבינציה של יחידות המדידה בצד השני של המשוואה (ראו Rakover, 2002, 2018 ). חוק הגישור אינו יכול לעמוד בדרישה זו. מדוע? משום שיחידות המדידה של המצבים, התהליכים הנוירופיזיולוגיים המתאימים שמופיעות בצד האחד של החוק, יחידות אלו שונות לחלוטין מיחידות המדידה המופיעות בצד השני של החוק והמשתייכות למצבים, לתהליכים המנטליים. (לאמתו של דבר, איש עדיין אינו יודע כיצד למדוד מצבים מנטליים באופן ישיר, אלא רק באופן עקיף על ידי פירוש של תצפיות בהתנהגות כולל דיווח וורבאלי.) קשה אם כן למצוא קנה מידה אחיד ומשותף לשני סוגי המדידות האלו, כך שמבחינה זו אין אנו דנים בחוק גישור המבוסס על זהות בין מושגים של שתי תיאוריות שונות, אלא לאמתו של דבר אנו מטפלים כאן במקרה הטוב באסוציאציה, בקורלציה (כמו למשל, הקורלציה מהסוג המנסה למצוא איזשהו קשר בין גודלן של עגבניות לבין מידת השימוש בשפה גבוהה בסוגים שונים של ספרים ועיתונים).

וכדוגמה אחרונה לדיון בסעיף זה, אראה שקשה מאוד אם לא בלתי אפשרי לתרגם הסבר מטרתי להסבר סיבתי. נסתכל על הדוגמה הבאה, 'אורי נהג במכוניתו מחיפה לתל-אביב *במטרה* לפגוש שם את חברתו יפה'. אפשר להציע תרגום זה של הסבר מטרתי (פעולה הנעשית כדי לממש *מטרה* מסוימת) להסבר סיבתי (הפעולה *נגרמה על ידי סיבה מסוימת*), 'הרצון לפגוש את יפה בתל-אביב *גרם* לאורי לנהג במכוניתו מחיפה לתל-אביב'. תרגום זה התבסס על הרעיון הפשוט של הפיכת המטרה לסיבת הפעולה. אולם כפי שיתברר מיד, תרגום זה מעורר בעיות גדולות המערערות על עצם הרעיון להעמיד הסבר מטרתי על הסבר סיבתי. ראשית, הרעיון של הפיכת המטרה לסיבה שגורמת לפעולה על ידי ייצוגה בהכרתו של האדם כגורם מנטאלי האחראי לפעולה (הנהיגה במכונית), למרות שרעיון זה נתפס במבט ראשון כטבעי ביותר, מעורר את בעיית הגוף/נפש: איך תהליך מנטאלי (מחשבה) מסוגלת לעורר התנהגות (וההפך איך התנהגות מעוררת תהליכים מנטאליים)? לבעיה זו, בעיית הגוף/נפש או בעיית המוח/הכרתיות אין עד היום שום פתרון המניח את הדעת. אם זה כך, הרי שעולה שהתרגום המוצע אינו פותר את בעיית ההסבר המטרתי, אלא רק מכניס אותה לבעיה חדשה. אם קודם לא הבנו כיצד מאורע עתידי מסוגל להסביר התרחשותו של מאורע בהווה או בעבר, הרי שעתה איננו מבינים כיצד מאורע מנטלי מחולל מאורע פיזיקאלי.

ולבסוף נקודה נוספת המעידה שהתרגום מהסבר מטרתי להסבר סיבתי מעורר בעיות קשות ביותר. מבחינה מתודולוגית הסיבה והתוצאה נתפסים כמאורעות שונים ונפרדים, כמו למשל, המאורע הסיבתי, מכה בעזרת מקל הביליארד, והמאורע התוצאתי, התגלגלות הכדור הלבן על שולחן הביליארד. הבחנה זו בלתי אפשרית במקרה של ההסבר המטרתי, כי קיימת תלות בין המושגים המופיעים בהסבר: הרצון של אורי, האמונה של אורי והפעולה של אורי. הסיבה היא הרצון של אורי לראות את יפה בתל אביב; התוצאה היא הפעולה של אורי, כלומר, הנסיעה מחיפה לתל-אביב שהינה מימוש האמונה של אורי שבעזרת הנסיעה הוא יגיע למטרתו – לפגוש את יפה בתל-אביב. במילים אחרות, הסיבה המנטלית לנסיעה מייצגת גם את הרצון של אורי וגם את המטרה משום שהרצון מכוון תמיד כלפי מטרה מסוימת; ובאותה מידה ניתן לומר שהאמונה שמה שיממש את הרצון מייצגת גם את הפעולה שתכליתה לממש את המטרה שאותה חפץ אורי להשיג; פעולת הנסיעה של אורי אינה סתם פעולה ללא מטרה וכיוון, זו פעולה בעלת משמעות המושרת על ידי הרצון והאמונה של אורי. כפי שניתן לראות משתי הדוגמאות האלו, זו של הביליארד וזו של אורי, בעוד שהסיבה והתוצאה בדוגמה של הביליארד הינם מאורעות נפרדים, הרי שבדומה של אורי הסיבה והתוצאה שזורים זה בזה ולא ניתן להפריד ביניהם.

1. *האם הצליחו לפתח מחשב מאוד מסובך ומתוחכם המסוגל כמונו להגיב (להימצא) בהכרתיות?*

באופן כללי ניתן לומר כי המחקר המטפל באינטליגנציה מלאכותית (Artificial Intelligence (AI)) לא הצליח לפתח מחשבים בעלי הכרתיות (ראו McDermott, 2007; Sun & Franklin, 2007). היות ו AI מושתתת על הגישה הפילוסופית הנקראת "פונקציונליזם" (functionalism) שמהווה את הבסיס לגישה של הפסיכולוגיה הקוגניטיבית, אתמקד בהקשר הנוכחי על גישה פילוסופית זו באופן ביקורתי.

גישת הפונקציונליזם מציעה פתרון חלופי לבעיית הגוף/נפש, פתרון שהוצע על ידי תאוריית הזהות בין MS לבין NS Neurophysiological State (NS))). לפי גישה זו מוגדר ה MS כמושג המחבר בין הגירוי לתגובה, שנמצא באינטראקציה עם MSs אחרים, ושניתן למימוש באופן חומרי (למשל, על ידי הנוירופיזיולוגיה של המוח). אחד החוקרים החשובים, Kim (1996), כותב בעניין זה:

"…psychological concepts are like concepts of artifacts. For example, the concept of 'engine' is silent on the actual mechanism that realizes it – whether it uses gasoline or electricity or steam … As long as a physical device is capable of performing a certain specified job, namely, that of transforming various forms of energy into mechanical force or motion, it counts as an engine. The concept of engine is specified by a job description of mechanisms that can execute the job." (p. 75)

ועתה עולה השאלה הבאה: אם ה MS מוגדר באופן פונקציונאלי כפי שתואר לעיל, האם יתכן שגם מצב מחשב CS (Computer State (CS)) יגלה הכרתיות בדיוק משום שהגדרת ה CS זהה לזו של ה MS? אם התשובה חיובית, הרי שנמצא הסבר מכניסטי להכרתיות. אולם הספרות מציעה מספר רב של טיעוני נגד. אחד הטיעונים החזקים והמפורסמים ביותר נגד תשובה חיובית לשאלתנו זו הוא הניסוי המחשבתי "החדר הסיני" של Searle (1980).

הבה נניח שהצליחו לבנות מחשב מאוד מורכב ומתוחכם המחקה את פעולתו של המוח האנושי. האם מחשב כזה יפתח הכרתיות כשל בני האדם? התשובה של הפונקציונליסטים, כאמור, תהיה חיובית. נניח שמחשב זה הצליח לענות על שאלות שהוצגו לו בשפה הסינית. האם המחשב המתוחכם מבין סינית כאדם שנולד בסין? שוב התשובה של הפונקציונליסטים חיובית, אולם לפי Searle (1980) התשובה היא שלילית. הרעיון היסודי העומד בבסיס הטיעון של Searle, טיעון החדר הסיני, הוא זה: נניח ש Searle עצמו, שאינו מבין מילה בסינית, לוקח על עצמו לבצע את כל האופרציות שהמחשב המתוחכם עשה על השאלות בסינית. Searle מקבל מצד אחד של החדר הסיני, המהווה מעין אנאלוגיה ליחידת החישובים של המחשב המתוחכם (מהצד האנאלוגי לקלט של המחשב), סידרה של סימנים סיניים, הולך אתם לספר הוראות באנגלית הקובע מה שעליו לעשות עם הסימנים האלה, מבצע את שהספר פוקד עליו לעשות, ומוציא מהצד השני של החדר הסיני (מהצד האנאלוגי לפלט של המחשב) סדרה חדשה של סימנים סיניים. מומחי לשון סיניים בוחנים את הפלט וקובעים חד משמעית שהפלט הינו תשובה אינטליגנטית בסינית לקלט, לשאלה בסינית שהוכנסה לחדר הסיני.

עובדה אחרונה זו אומרת ש Searle עבר את המבחן המפורסם של Turing (1950) . לפי מבחן זה, אם נותנים למחשב ולאדם סידרת שאלות, ומספר שופטים אנושיים אינם מסוגלים להחליט על סמך התשובות מי האדם ומי המחשב, אזי אין לשופטים שום הצדקה לשלול מהמחשב את תכונת ההכרתיות. השופטים הסיניים קבעו שהתשובות שיצאו מהחדר הסיני מעידות על כך שהמשיב הוא אדם סיני אינטליגנטי, אולם Searle שיצא מהחדר הודיע קבל עם ועדה שהוא אינו מבין מילה אחת בסינית. הוא אמנם ביצע את כל הפעולות התחביריות הסיניות הנדרשות לעשות על הסימנים הסיניים, אך נשאר בחוסר הבנה מוחלט של שפה זו. ומכיוון ש Searle ביצע בדיוק את כל הפעולות של המחשב המתוחכם ולמרות זאת הוא נשאר חסר הבנה בסינית, עולה שגם המחשב המתוחכם אינו מבין את אשר הוא ביצע בסינית. כלומר, ההשערה של הפונקציונליזם על אודות ההכרתיות הופרכה! (כאן עלי להעיר הערה על האפשרות הבאה. אפשר להעלות השערה שאולי מחשבים קוונטיים ולא דיגיטליים עשויים לפתח הכרתיות בגלל היכולת האדירה שלהם. למרות שבמקצועי איני מתמטיקאי ולא מדען מחשבים, אני מטיל ספק גדול באפשרות זו בגלל שקראתי שכל בעיה שמחשב קוונטי מסוגל לפתור, יפתור גם מחשב דיגיטלי, אבל במשך זמן ארוך פי כמה וכמה מהזמן הנחוץ למחשב הקוונטי.) אסיים דיון ספציפי זה בהערה הבאה: הניסוי המחשבתי של Searle עורר וויכוחים רבים בספרות המקצועית, שלא אדון בהם כאן, אך אציין שפרסמתי בזמנו מאמר התומך בגישתו של Searle המבוסס על הרעיון שלפיו מחשב-על חיקה את Searle בחדר הסיני (ראו (Rakover, 1999. ניתוח כל האפשרויות העולות מהשאלה האם מחשב על זה פתח הכרתיות, העלה תשובה שלילית.

לסיום הטיפול בשאלת הסעיף הנוכחי, אדון בשאלה הבאה: האם האנאלוגיה בין פעולתו של המחשב המבוססת על הקשר בין התוכנה לחומרה מצליחה להסביר את היחס בין ההכרתיות לבין המוח? אומר מיד שהתשובה לשאלה זו הינה שלילית. אך כדי להבהיר אותה עלי להסביר בקיצור נמרץ את הדרך שבה עובד המחשב הדיגיטלי. ברגע שאנו כותבים משהו על המחשב כמעבד תמלילים, עובר המשפט המובע בשפה האנושית (למשל, עברית, אנגלית) מספר תרגומים שבסופם הופך הכתוב לסדרות ארוכות של אפסים ויחידות, כשאלה מזוהים במחשב למצבים של יש זרם חשמלי = 1, ואין זרם חשמלי = 0. כלומר, כל מה שאנו מעוניינים שהמחשב יבצע כמו למשל, חישובים מספריים, חישובים בעזרת סימבולים, כתיבה בשפה, ציור וכו עוברים בעזרת תוכנות מסוימות לסדרות ארוכות של אפסים ויחידות, כשמה שהמחשב עושה פיזית הוא בעזרת שני מצבים יש-אין חשמל המתורגמים בחזרה לשפה שבה מעוניין האדם (see, e.g., Block, 1995; Deitel & Deitel, 1985; Von Eckardt, 1993).

מספר חוקרים סבורים שהיחס בין תהליך הכרתי למוח באדם אנאלוגי ליחס בין תוכנה לחומרה במחשב ומדגישים שכמו שכתיבה בשפה אנושית בעזרת המקלדת של המחשב מתפרקת בסופו של דבר לסדרות של מצבי חשמל, כך מתפרק תהליך הכרתי באדם לנוירופיזיולוגיה של המוח. למשל, Dennett (1979) מתאר תהליך פירוק זה בעזרת top-down strtegyממושגים מנטליים למרכיבים ביולוגיים:

… a top-down strategy that begins with a more abstract decomposition of the highest levels of psychological organization and hopes to analyze these into more and more detailed smaller systems or processes until finally one arrives at elements familiar to the biologists." (p. 110)

הוא ממשיך ומצייר את תהליך הפירוק בצורה הבאה: תהליך מנטלי מיוצג בעזרת תרשים זרימה הכולל קופסאות קטנות שקשורות זו לזו על ידי חיצי זרימה כשבכל קופסה נמצא איש-קטן (homunculus) הממלא תפקודים מסוימים:

"If we then look closer at the individual boxes we see that the function of each is accomplished by subdividing it via another flow chart into smaller, more stupid homunculi. Eventually this nesting of boxes within boxes lands you with homunculi so stupid (all they have to do is remember whether to say yes or no when asked) that they can be, as one says, 'replaced by machine'." (p. 124)

האם תפיסה זו שלפיה ההכרה-מוח פועלים באנאלוגיה לתוכנה-חומרה במחשב, אכן נכונה? כלומר, האם אכן ניתן לפרק את ה MS, התהליך המנטלי, ליחידות כה אלמנטריות שניתן לבסוף לזהות אותן עם יחידות אלמנטריות נוירופיזיולוגיות, ה NS הבסיסיים? לדעתי התשובה שלילית. בספרות המקצועית מופיעים מספר נימוקים לכך ואני אדגיש רק את הבאים (ראו למשל אצל Rakover, 2018):

(א) האנאלוגיה הנ"ל בין המחשב למוח האדם והכרתו אינה לוקחת בחשבון שהמובן והמשמעות של סידרת הסימנים שהמחשב מדפיס או שמופיעות על מסך המחשב נמצאים רק בתודעתו של בני אנוש, אלה המשתמשים במחשב ואלה שטרחו ובנו את המחשב כדי שימלא פונקציות מסוימות לשימושו של האדם. ללא עין אדם הסימנים הפיזיים בין כשהם נכנסים למחשב דרך הקלט או בין שהם מתורגמים למצבים חשמליים בגוף המחשב ובין שבסופו של דבר הם נפלטים כסדרה של סימנים כתוצאות של עיבודן הפיזי, והסמנים האלה, בכל השלבים, אין להם לסימנים הפיזיים, למאורעות הפיזיים, שום משמעות ומובן ללא תודעתו של האדם. אם למשל, יתגלה בחפירות ארכיאולוגיות לוח מלא סימנים מוזרים ונשער שלוח זה מראה כתב של תרבות עתיקה וקדומה מאוד, לא נוכל להבין את הכתוב כי משמעות הסימנים לא הועברה מדור לדור עד לימינו אנו.

(ב) בעוד שהתרגום של סידרת הסימנים משלב לשלב במחשב אינו גורע אינפורמציה ואינו מוסיף אינפורמציה, אצל בן האנוש אין הדבר כך. כדי להמחיש עניין זה נבחן כדוגמה פשוטה, את שלושת המשפטים הבאים:

1. רות אוהבת את החתול

2. רות אוהבת כריך ביצה

3. רות אוהבת לרקוד

שלושה משפטים פשוטים שלא יקשה בעד המחשב לפרק אותם תחילה למילים בודדות ולהרכיב אותן לאחר מכן בהתאם למבנה המשפט הפשוט: נושא, פועל, ונשוא. הנקודה החשובה שאני מעוניין להדגיש כאן היא שמשמעות המילה 'אוהבת' שונה לחלוטין בשלושת המשפטים האלה, כי מובנה מוקנה לה דרך המשפט בכללותו. המחשב עיוור למשמעות ולמובן זה, עבורו המילה 'אוהבת' היא סתם סימן פיזי שיש להכניסו במקום הנכון לפי כללי תחביר מסוימים הקשורים במבנה המשפט. למה הדבר דומה? להבנתו (במרכאות כפולות ומשולשות) המדהימה של המחשב שלי שבאה לידי ביטוי במקרה הבא: יום אחד כתבתי מאמר ובו כתבתי שיום אמר כך וכך (הכוונה הייתה לפילוסוף David Hume) והינה המחשב הזריז והחכם שלי מיד קפץ והודיעה לי באותיות של קידוש לבנה את השעה, את היום, ואת השנה המתאימה לכתיבת המאמר (בעברית למילה 'יום' יש משמעות שכיחה של Day).

(ג) בעוד שמושג האינפורמציה מוגדר היטב במדעים, בפסיכולוגיה מושג זה פרוץ לרווחה והוא משמש כתיאור לאוסף של מושגים, תהליכים כמו למשל, תוכן המודעות, משפטים, מילים, הברות, תהליכי זיכרון, תפיסה, הלוצינציות ... ומה לא (e.g., Palmer & Kimchi, 1986). ברור שבמצב כזה כשהמושג החשוב ביותר בפסיכולוגיה הקוגניטיבית אינו מתוחם אפילו בצורה הרחבה ביותר, קשה אם לא בלתי אפשרי להעמיד את מושג ההכרתיות על מושג האינפורמציה.

(ד) כפי שתואר לעיל על ידי הניסוי המחשבתי על מרי מדענית הראייה, שלמדה משהו חדש כאשר ראתה בפעם הראשונה בחייה צבע אדום (על אף שידעה את כל שאפשר לדעת על ראיית צבעים מבחינה פיזיקאלית ופיזיולוגית), ההסבר השלם של תפיסת צבעים יהיה חסר ולכן ניתן לומר שהסבר מסוג זה יסבול מפער הסברי (explanatory gap, see Levine, 1983). ושוב עולה מכאן שתפיסת היחס בין הכרה לבין מוח כיחס בין התוכנה לבין החומרה במחשב אינה נכונה.

טבלה 4.1 מסכמת בקיצור את המסקנות העיקריות של הסקירה הנוכחית:

**============================================================**

**טרם נמצא הסבר להכרתיות במושגים המקובלים במדעים ובפסיכולוגיה**

1. **קשה להסביר התנהגות באופן שלם מבלי להשתמש במושג ההכרתיות**
2. **טרם הצליחו לפתח תיאוריה נוירופיזיולוגית של המוח המסבירה הכרתיות**
3. **טרם הצליחו להעמיד הסבר המעוגן בהכרתיות על הסבר נוירופיזיולוגי**
4. **עדיין לא פתחו מחשב מסובך ומתוחכם המגלה הכרתיות**

**=========================================================**

**טבלה 4.1 מציגה את המסקנות העיקריות מהסקירה הספרותית על האפשרות להסביר הכרתיות באופן מכניסטי**

**(3) הכרתיות כמושג מסביר אך לא מוסבר**

הכרתיות היא מושג המתייחס לתופעה מנטאלית-התנהגותית המופיעה בבני אדם (וגם בבעלי חיים עילאיים), ושניתן לתארה כעולמו הפנימי של היחיד שרק הוא מסוגל לחוש ולצפות בו. נראה שעד כה אין הסבר המניח את הדעת לא לתופעת ההכרתית ולא לבעיית היחס בין ההכרתיות לבין המוח, כלומר, טרם נמצאה תיאוריה שתסביר יפה את הקשר בין הגוף לנפש, בין המוח לבין ההכרתיות. מתברר שכל תיאוריה אשר ניסתה להעמיד את ההכרתיות על הנוירופיזיולוגיה של המוח, לא הניבה פירות טובים. תיאוריה זו, בין אם זו תיאורית הזהות, הפונקציונליזם, ריבוי מימושים או גישת הפיקוח-התערבות (supervenience) שלפיה לא יתכן שינוי ב MS ללא שינוי ב NS, לא עלתה יפה, בכל מקרה הועלו בעיות חמורות נגד התיאוריה המוצעת והיא נדחתה על ידי רוב החוקרים.

לדעתי, אפשר לסכם את בעיית ההסבר של ההכרתיות על ידי תהליכים נוירופיזיולוגיים במוח בצורה שמעלה את הבעיה הבאה: מצד אחד, הניסיון להעמיד את ה MS על ה NS הוא מפתה מאוד, כיוון שאז אפשר להסביר את ההתנהגות בכללותה (כולל ההתנהגות ההכרתית) על ידי הסתמכות על התיאוריות הסיבתיות הנוירופיזיולוגיות הידועות במדע (אולם כדאי לשים לב להשלכות המוזרות העולות מתיאוריה כזו המעמידה את ההכרתיות על הנוירופיזיולוגיה של המוח – ראו לעיל). מהצד האחר, אם ההסבר להתנהגות מטופל כולו על ידי ה NS, מהו ערך ההסבר וחשיבותו של ה MS? עבור מה אנו זקוקים לו? שאלות אלו מנוגדות לשכל הישר, כלומר, להתנסות היומיומית ולמודל ה CM המסבירים את התנהגותו של היחיד על ידי הסתמכות על עולמו הפנימי של האדם, על ההכרתיות (למשל, הלכתי לקולנוע משום *שרציתי* לראות את השחקנית גל גדות בסרט "wonder-woman"). (דיון רחב ביכולת ההסבר של ההכרתיות ושל משמעויות-החיים ראו בהמשך הפרקים הבאים.) אם הכול מוסבר בעזרת תהליכים נוירופיזיולוגיים, על ידי ה NS, יוצא של MS אין כל ערך הסברי ולכן כל התיאוריות המבוססות-נוירופיזיולוגית מביאות אותנו כמעט בהכרח לגישה של האפיפנומנליזם (epiphenomenalism) – הגישה הישנה שנדחתה זה מכבר, ושלפיה ל MS אין כל כוח הסברי. וכך כותב Kim (2002) בעניין זה בתקציר לספרו מ 1998:

"To summarize, then, the problem of mental causation is solvable for cognitive/intentional mental properties. But it is not solvable for the qualitative or phenomenal characters of conscious-experience. We are therefore left without an explanation of how qualia can be causally efficacious; perhaps, we must learn to live with qualia epiphenomenalism." (p. 643)

אינני מקבל את התפיסה של Kim (2002) שלפיה ההכרתיות (conscious-experience) הוא בחזקת אפיפנומנה. אני מציע גישה הפוכה, מנוגדת, אפיפנומנליזם-הפוך (anti or reverse- epiphenomenalism) שלפיה לא רק שלהכרתיות יש השפעה ולכן אינה אפיפנומנה, אלא שיש לראותה כמושג מסביר ראשוני, בדיוק משום שעדיין לא נמצא לו הסבר המניח את הדעת. כלומר, לאור הפרק הנוכחי שלפיו טרם נמצאה תיאוריה שמסבירה את ההכרתיות על בסיס הנוירופיזיולוגיה של המוח, יש מקום להצעה הבאה: הבה ונראה בהכרתיות מושג תיאורטי-מסביר ראשוני שאינו ניתן להסבר על ידי מושגים בסיסיים יותר. הצעה זו דורשת את ההבהרות הבאות.

ראשית, איני מציע שבגלל שטרם נמצא הסבר לבעיית ההכרתיות, מן הראוי להניח שהכרתיות אצל בעלי חיים, בפרט האדם, היא בחזקת כוח טבע חדש. הנחה זו עשויה ליצור בלבול עצום בתשתית ההסבר המכניסטית המקובלת (למשל, חוקי שימור האנרגיה עלולים להישבר) (ראה טיעונים דומים אצל Carroll, 2016).

שנית, איני טוען שההכרתיות אינה תלויה בתהליכים מוחיים באדם ובבעלי חיים, אלא שם דגש על כך שטרם נמצאה תיאוריה המסבירה את הקשר בין שני אלה. כתוצאה, אני מניח רק שהכרתיות הוא מושג מסביר אך לא מוסבר, כלומר שהוא מושג מסביר ראשוני וניתן לאפיין אותו כמתייחס למצב של reverse- epiphenomenalism.

בינתיים, אם כן, אניח שהתכונות החשובות של ההכרתיות הן אלו:

(א) ההכרתיות מצויה בדרגות שונות אצל כל אדם ואדם (וגם אצל יצורים חיים אחרים);

(ב) רק האדם עצמו חש באופן הכרתי את תוכן הייצוגים השונים המופיעים בראשו (ה MSs);

(ג) ללא הכרתיות היה האדם מתפקד אך ורק ברמה הפיזיולוגית והיה נחשב לאמתו של דבר כמי שנמצא במצב הדומה לזה של צמח, מעין זומבי-פילוסופי (יצור דמיוני המתנהג כאדם אך משולל כל הכרתיות);

(ד) ההכרתיות מסוגלת להשפיע על תפקודו הפיזי של האדם;

(ה) ההכרתיות עשויה להשתנות כתוצאה מהשפעות פיזיקאליות, למשל, גירויים תחושתיים

כמו אור וצליל שמעוררים ביחיד תחושות הכרתיות האופייניות לגירויים אלה, ראיה ושמיעה;

(ו) ההכרתיות תלויה בתפקוד הנורמאלי של המוח;

1. ההכרתיות היא שמאפשרת כתנאי הכרחי את התחושות האופייניות לגירויים

תחושתיים, את משמעות-החיים ותחושת היות-חי, ואת ההבנה.

**פרק (5) משמעות-חיים, אבסורד, והתאבדות (I): מנקודת ראות של הניתן להסבר**

אפשר להתייחס למושגים הבסיסיים ביותר בגישה האקזיסטנציאליסטית כמו משמעות-חיים, אבסורד, והתאבדות, כמושגים המתארים סוגים מסוימים של התנהגות (למשל, התנהגות קוגניטיבית, אמוציונאלית, מנטליסטית) וככאלה ניתן לנסות ולהסביר אותן. הפרק הנוכחי מתאר כיצד מטפל מודל ה CM במושגים אלה, באיזה תנאים משמעות I נחלשת ומתי היא ממשיכה לתמוך באדם למרות שהוא נקלע למשבר חיים חמור הפוגע בדרך כלל במשמעות II. כדוגמה, מנתח הפרק את פרשת משבר החיים של הסופר הרוסי הגדול, טולסטוי.

שלושת המושגים האלה נחשבים למושגי יסוד בפילוסופיה האקזיסטנציאליסטית. למשל, ב*המיתוס של סיזיפוס*, קובע קאמי (1992):"בעיה פילוסופית רצינית באמת יש רק אחת: ההתאבדות." (עמ' 13). ללא כל ספק שמושגים אלה ניתנים לדיון מנקודת הראות של הפילוסופיה. מדוע התאבדות? כי זו תגובה לחוסר המשמעות בחיים, לכך שהחיים הינם אבסורד משום שאנו חפצים להסביר עולם שלא ניתן להבנה ומשום שבין כה וכה סופנו הוא מוות. ברגע שאנו מתחילים לתהות מהו טבעו של מושג זה 'משמעות-חיים', מהי מהותו של מושג זה? האם, למשל, משמעות-חיים מצויה בעולם עצמו או שמשמעות-חיים היא יצירתו של האדם (ואם כן, האם גם בעלי חיים עילאיים ניחנים במשמעות-חיים בדרגה כל שהיא?), האם לכל אדם יש משמעות-חיים אינדיבידואלית משלו? או אולי משמעות-חיים מוקנית לנו על ידי החברה? או שהיא ניתנת לאדם מיד עם הלידה והיא מתפתחת עם הזמן, כפי שמתפתחת יכולת הלמידה של האדם ושל בעלי החיים? האם ניתן להצדיק וכיצד ראוי להצדיק דרך חיים המהווה משמעות-חיים לאדם מסוים (אך לאו דווקא לאדם אחר)? באותו הרגע שאנו מעלים שאלות מהסוג הזה, אנו מתייחסים אל המושג 'משמעות-חיים' מנקודת ראותה של הפילוסופיה. התשובה שאני נתתי לשאלות אלה (ראה לעיל) אומרת שעצם היות האדם במצב של הכרתיות מקנה לו משמעות-חיים בסיסית (משמעות-חיים I) והיא זו שנותנת לחיים את טעמם היסודי, דהיינו, שבלי הכרתיות נלקחת מהאדם משמעות חייו כפי שנלקחת ממנו היכולת להבנה, כלומר, שללא הכרתיות אין משמעות-חיים ואין הבנה. תשובתי זו שייכת מצד אחד לתחום הדיון הפילוסופי, אולם מהצד האחר, לתשובה זו יש השלכות אמפיריות ולכן אפשר לדון בה גם מנקודת הראות המדעית-תיאורטית.

בספרו על *אלבר קאמי והפילוסופיה של האבסורד,* דן שגיא (2000) במושג 'האבסורד' מכמה נקודות ראות. לדוגמה, האם, לפי קאמי, מהווה מושג האבסורד אבן פינה, יסוד, בפילוסופיה של הקיום האנושי? אפשר לתת תשובה חיובית לשאלה זו בהסתמך על שני הספרים של קאמי, *הזר*,ו*המיתוס של סיזיפוס.* לפי הספר האחרון עולה שהאבסורד שבא לידי ביטוי בחייו של סיזיפוס הוא תחום החיים של האדם – אנחנו חיים בעולם שמעצם מהותו הוא אבסורדי. זהו הנתון הבסיסי של חיי האדם. וכמו שציינתי זאת לעיל, מעשיו של מֶרסו, היחס למות אמו והרצח שהוא בצע, אינם מתיישבים עם ההיגיון, הם אבסורדיים; וכך גם התנהלותו של המשפט של מרסו באשמת רצח – התנהלות משונה, אבסורדית. אולם ברומן *הדבר* ובמסה הפילוסופית *האדם המורד* תופס קאמיאת מושג האבסורד בצורה שונה. אין זה מושג בסיסי, אבן ראשה, המתאר את מצבו הקיומי של האדם – חיים באבסורד (במרד), אלא נקודת פתיחה להבנת מצבו של האדם. בדומה לשיטת הטלת הספק של דקרט, שבעזרתה מבסס הפילוסוף הצרפתי את הידע בנקודת וודאות, משתמש קאמי באבסורד כדי להגיע בסופו של דבר לתפיסת האדם כמרכיב בחברה, ולהבנת האחריות והסולידאריות בין בני האדם. וכך כותב קאמי (1971) בעניין האבסורד: "אלא שנבצר מסתירה יסודית זו שלא להתבלט בהמונן של סתירות אחרות, משעה שאנו מתיימרים להתקיים באבסורד, תוך התעלמות מאפיו האמתי, שהוא – להיות חוויה שנחותה, נקודת מוצא, שווה-הערך הקיומי של הספק המתודי לדקרט. האבסורד הוא סתירה בתוך עצמו." (עמ' 11).

נקודת התייחסות נוספת אל האבסורד מתבטאת בשאלה הבאה: האם, לפי קאמי, מהווה האבסורד התנסות ראשונית, תחושתית של העולם האבסורדי? התשובה החיובית מפרטת אופנים שונים של חווית האבסורד, למשל, המציאות האבסורדית נתפסת כמורכבת ממאורעות עוקבים שאין ביניהם שום קשר הגיוני, ללא כל משמעות מאחדת; היות והמציאות העכשווית נתפסת כחסרת משמעות, בלתי מובנת, הופך העולם לזר, מנוכר, מפחיד ומאיים, וההווה הופך להיות חסר ערך כשהאדם חי בצפייה לעתיד לא ברור לחלוטין; מה מחכה לנו בעתיד? רק דבר אחד וודאי – המוות שהופך להיות מרכיב חשוב ביצירת האבסורד ובתחושה שאין שום הגיון בחיים ומוטב לסיים את הפארסה חסרת-ההיגיון הזו בהתאבדות. קודם לכן ציינתי שמהאבסורד אפשר להסיק לאמתו של דבר כל מסקנה אפשרית: התאבדות, רצח, הוללות, ופרישות, משום שמהנחות היוצרות סתירה עצמית ניתן להסיק כל דבר. עם זאת, כדאי לציין שאפשר להציע למשמעות-חיים הגדרה כזו, שממנה תנבע ההתאבדות באופן טבעי וחלק. למשל, אם נגדיר משמעות-חיים כטעם, הסיבה שבגללה יש להמשיך לחיות על אף כל הסבל הכרוך בכך, אזי אובדן משמעות-החיים (כתוצאה מהיות החיים אבסורד) יוביל באופן ברור לחיסולה של הסיבה להמשיך לחיות, ולכן עלול היחיד להעלות בדעתו שמן הראוי לסיים את האבסורד הזה של החיים אחת ולתמיד.

כפי שניתן לראות מהניתוח הקצר לעיל, קאמי מתייחס אל מושגי האבסורד ומשמעות-חיים, משני נקודות ראות: האחת, כמושגים פילוסופיים המעידים על מהות האדם ועולמו; השנייה, כהתנהגויות אנושיות הדורשות פרשנות שנסובה סביב המושג היסודי 'אבסורד'. נדמה שאם קאמי עצמו לא היה חש בהכרתו ובליבו קיום אבסורדי או אם לא היה רואה בהתנהגותם של אחרים עדות המעידה על אבסורד, לא היה מפתח פילוסופיה של אבסורד. ניתן אם כן להציע שהתנהגות אנושית מוזרה, בלתי מובנת, אבסורדית היא תנאי הכרחי לפילוסופית האבסורד שאותה פיתח קאמי. נקודת ראות התנהגותית זו עלולה להביא חוקרים לידי המחשבה שאולי הטיפול באנשים הסובלים מאובדן משמעות-חיים, מניכור ופחד מפני עולם אבסורדי, הוא עניין פסיכולוגי ולא פילוסופי. אפשר גם לתמוך בגישה פסיכולוגית זו ולומר בעקבות הצעתו של איגלטון (ראה לעיל) שהתפיסות הפילוסופיות של קאמי ושל סארטר נוצרו בתקופות של משבר איומות, מלחמת העולם הראשונה והשנייה. ייתכן אם כן שאפשר לראות בתפיסותיהם של שני פילוסופים אלה תגובות המבטאות את המשבר האינטלקטואלי החמור שמוראות המלחמה הציגו בפני העולם: אבסורד המתבטא במעשי רצח נתעבים הנעשים בידי האדם שמצדיק את מעשיו המאוסים והמשוקצים בעזרת אידיאולוגיות מגונות ומבעיתות כמו הפשיזם שטיפח גזענות ארית והצדיק את השמדת היהודים כטפילים המזהמים את הגזע הארי הטהור והעליון.

נגד תפיסה זו, המנסה להסיט את מרכז הדיון במושגים הנידונים לכיוון פסיכולוגי, מנסה Landau (2017) להעלות מספר טיעונים התומכים בניתוח הפילוסופי של מושגים אלה. וכך הוא כותב:

"Some might claim that meaning and meaningless are solely a matter of moods, emotions, and other psychological states. Hence, philosophical and rational discussions such as those presented in this book are irrelevant. According to this view, issues relating to the meaning of life should be dealt with by psychologists, …" (p. 254).

Landau (2017) אינו מסכים לתפיסה זו ומביא מספר שיקולים ודוגמאות המראות ששימוש בארגומנטים לוגיים, בעיונים וברורים פילוסופיים, עשויים להיות מאוד רלוונטים ולשנות את דעתו הפסימית של האדם שאיבד את דרכו בחיים לכיוון האופטימי. העובדה שלרגשות יש מרכיבים קוגניטיביים-הגיוניים ברורים עשויה להיות בסיס לשינוי דעתו של האדם שמשמעות חייו נגוזה בעזרת טכניקות פילוסופיות.

אין לי כל ספק שבדברי Landau (2017) מצויה אמת רבה (הרי במקרים רבים אפשר לשכנע בני אדם לשנות את דעתם בדברי טעם), ואין לי כל צל של ספק שניתן לטפל במשמעות-חיים ואבסורד בכלים פילוסופיים (והרי כך אפשר לקרוא את הדיון שקאמי וסארטר הציגו בכתיבה התיאורטית והספרותית שלהם). אולם בפרק הנוכחי אני מעוניין לדון במושגים הנידונים בעזרת השאלה: כיצד ניתן להסבירם וכיצד ניתן להשתמש במודל ה CM על מנת להסביר תופעות התנהגותיות אלו? שאלה זו מובילה לשאלה הרחבה יותר, השאלה הכללית הבאה.

**בהתייחס למושגים הנידונים, מהו המוסבר והמסביר ומהי הפרוצדורה שבעזרתה מסבירים?**

כדי לענות על השאלה הנוכחית, נבחן תחילה את הדרכים למתן הסבר בדיסציפלינות ובמצבים שונים.

1. **מתמטיקה**:במתמטיקה מסבירים באופן תיאורטי השערות, משפטים מתמטיים על ידי כך שמספקים *הוכחה* להשערה מסוימת, למשפט, כלומר, מוציאים מסקנה מהנחות נתונות על ידי הסקתה בעזרת כללים לוגיים-מתמטיים מסוימים.
2. **נורמות חברתיות:** במקרים של התנהגות נורמטיבית מתנהג היחיד בהתאם לכללים המקובלים. למשל, כשהוא עוצר באור אדום, מוסברת התנהגותו זו בכך שאומרים שהיחיד מבצע, הולך בעקבות הכללים הידועים לו ולכן הוא עצר את נסיעת מכוניתו כשהרמזור הפך מירוק לאדום.
3. **פילוסופיה:** בפילוסופיה ההסבר לובש צורה של *הנהרה, הבהרה, הצדקה, ומתן פירוש* למושגים ולתופעות כלליות בעולם ולתהליכים קוגניטיביים-הכרתיים האופייניים לכל אדם. פעולות אלה נעשות על ידי טכניקות שונות הכוללות בין היתר שימוש בדוגמאות אמפיריות הברורות לכולם, ניסויי מחשבה, בניית תיאוריות (מטאפיזיות בעיקר) ושימוש במטפורות המאירות את הנושא הנידון ומסירות מעליו ערפל.
4. **פסיכולוגיה:** בפסיכולוגיה *מסבירים* באופן אמפירי תופעות התנהגותיות בעזרת שימוש בתיאוריות וכללי הסבר, מודלים של הסבר שברוב המקרים שאולים מהמדעים.

ברור מאפיון קצר זה, שקשה מאוד לטפל בנושאי הספר הנוכחי בכלים מתמטיים או בעזרת נורמות חברתיות המקובלות על כולם (כמו למשל כללי התחבורה). מה שכן ניתן לעשות הוא לטפל במושגים אלה בגישה הפילוסופית, כפי שעשו קאמי וסארטר (ומה שנעשה עד כאן בחלק מהעבודה הנוכחית) או לטפל במושגים אלה כבמושגים התנהגותיים המתארים תופעות התנהגותיות, שלהן אפשר להציע הסברים – וזו לאמתו של דבר המטרה העיקרית של הפרק הנוכחי.

כדי לממש מטרה זו יש לציין מיד שכל אחד מהמושגים ההתנהגותיים האלה הוא בעל פנים כפולות, ולכן יש להסביר מדוע קבוצה של אנשים מתנהגת בהתאם למושג וקבוצה אחרת בניגוד לו. למשל, מצד אחד, יש להסביר מדוע קבוצה גדולה של אנשים מוצאים משמעות-חיים בדרגות שונות, אך מהצד האחר, יש להסביר מדוע מעטים מתייאשים מהחיים ולא מוצאים בהם שום משמעות. ובאותו אופן יש להסביר מדוע קבוצה גדולה מאוד מוצאת שהחיים יפים ולא אבסורדיים ואילו קבוצה אחרת, קטנה בהרבה, תופסת את העולם כאבסורד. אם, בהתאם לקאמי, ניתן לראות בהתאבדות אינדקס התנהגותי (קוגניטיבי-אמוציונאלי) לחוסר משמעות-חיים ותפיסתם כאבסורד, אזי הסטטיסטיקה הבאה עשויה לומר משהו חשוב על המושגים האקזיסטנציאליסטיים הנידונים. מתברר שמספר המתאבדים בעולם הוא אכן נמוך ביותר, הממוצע העולמי עומד על כ 15 מתאבדים מתוך 100.000 בני אדם (דהיינו, 0.015 אחוז), כאשר שיעור זה משתנה לפי מין האדם, גילו וארץ מוצאו (ראו, WHO: World Health Organization, suicide data). יש אם כן להסביר במקרה הנוכחי מדוע רוב רובם של בני האדם נאחזים בחיים על אף הייסורים הקשים ואילו מיעוט קטן מתאבד. כדי להציע הסבר מתאים אשתמש במודל ההכרתיות-משמעות (The Consciousness-Meaning (CM) model) שתיארתי אותו קודם לכן, ועתה לשם נוחות הקריאה וריענון הזיכרון אחזור ואסכם אותו בקיצור.

*ה CM model*מניח שאת המשמעויות לגווניהן ולדרגותיהן מייחס האדם לעולם (כלומר, לבני אנוש, לבעלי החיים, לצמחים ולדומם). המודל מבחין בין שלוש סוגים של משמעויות: *משמעות I-מולדת, משמעות II-נרכשת רגילה* (*משמעות II-רגילה*) ו*משמעות II-נרכשת קיצונית* (*משמעות II-קיצונית*). תחילה אתאר את משמעות I-מולדת. המודל מניח כהנחה בסיסית שבני אדם (ובמידה מסוימת בעלי החיים העילאיים) נולדים עם פעולת ייחוס אוטומטית של משמעות לעולם, כשמשמעות זו מושרת על ידי עצם החוויה ההכרתית הנוצרת על ידי *גירויים תחושתיים*, כגון, ראיה, שמיעה, ריח, טעם, מישוש, תנוחת הגוף (proprioception), תנועה, לחץ, חום, קור, כאב, ועונג (מגע, אוכל, מין). ההכרתיות מהווה איפא את משמעותן המולדת והבסיסית של החיים. אלו הן חוויות הכרתיות שדרכן חש האדם שהוא חי. לפי מודל ה CM ההכרתיות מהווה תנאי הכרחי למשמעות (כי כדי שאדם ירגיש, למשל, עונג מיחסי אהבה צרכים להתממש מספר גורמים נוספים שהקורא יכול לחשוב עליהם בעצמו).

על אף שאני מניח שהכרתיות *לגירויים תחושתיים* בתנאים נורמאליים משרה אוטומטית משמעות-חיים של היות-חי, תנאי זה מוגבל בתחום מסוים של *גירויים תחושתיים*. שינויים קיצוניים בדרגות ההכרתיות מלווה בשינוי בדרגות ואיכות המשמעויות, ממשמעויות נעימות וחיוביות למשמעויות בלתי נעימות, בלתי נסבלות, שליליות.

ועתה אעבור לתיאור של *משמעות II-רגילה,* ו*משמעות II-קיצונית*. האדם מסוגל ליצור בראשו או לקבל ולהפנים מטרות וכוונות שונות, ערכים, רעיונות ומחשבות, כלומר, *מאורעות מופשטים*, שעליהם מרעיפה הכרתיות תחושה של משמעות – *משמעות II*-*נרכשת (רגילה, קיצונית)*. ברוב המקרים, עוצמת המשמעות של *המאורעות המופשטים* היא נמוכה, אך היא עשויה להפוך לדרך חיים, למשמעות-חיים גדולה וחזקה בשני המצבים החשובים הבאים: (א) כשתוכן המאורעות המופשטים הינו בעל משמעות נרכשת הנוגעת ישירות לחיי היחיד עצמו, להישרדותו; (ב) כשהרעיונות המופשטים עוברים העצמה אמוציונאלית במשך זמן ארוך (מהילדות ואילך), ובמיוחד כאשר הם עוברים תהליך של הנחלת דעות, [עמדות](https://he.wikipedia.org/wiki/%D7%A2%D7%9E%D7%93%D7%94_(%D7%A4%D7%A1%D7%99%D7%9B%D7%95%D7%9C%D7%95%D7%92%D7%99%D7%94)), [אמונות](https://he.wikipedia.org/wiki/%D7%90%D7%9E%D7%95%D7%A0%D7%94) או [השקפת עולם](https://he.wikipedia.org/w/index.php?title=%D7%94%D7%A9%D7%A7%D7%A4%D7%AA_%D7%A2%D7%95%D7%9C%D7%9D&action=edit&redlink=1) בצורה אינטנסיביות קיצוניות, כלומר, כשהיחיד עובר תהליך של אינדוקטרינציה, למשל, דתית, אידיאולוגית, פוליטית.

ניתן להציע שמשמעות II-קיצונית, כמו אמונה דתית ואידיאולוגיה, עשויה להיעלם מחייו של אדם בשעות משבר גדולות, אולם משמעות I-מולדת אינה נפגעת ונשארת עם היחיד (אלא במקרים קיצוניים ביותר כמו מחלות קשות ואובדן בן משפחה קרוב). משמעות II-קיצונית עלולה להיהרס בשעת משבר, כמו במקרים של התפקחות מאידיאולוגיה או המרה דתית.יש עם זאת להדגיש שהעובדה שאנשים מסוגלים לעבור מדרך חיים אחת לשנייה, למשל, מפשיזם לדמוקרטיה, מאמונה באל לחילוניות, מראה שמשמעות I-מולדת נשמרת ברמה סבירה המאפשרת ליחיד להמשיך לחיות ולחפש דרך חיים חדשה.

בהינתן סיכום קצר זה של מודל ה CM, אעבור עתה לטפל בשאלה הבסיסית הנ"ל: איך ניתן להסביר שלרוב רובם של בני האדם יש משמעות-חיים, דרך חיים שלאורה הם הולכים? הסיבה היסודית היא שמשמעות I-מולדת נתונה לאדם מהרגע שהוא נולד ומשום שהיא מולדת לרוב רובם של בני האדם השאלה של קאמי על אודות ההתאבדות אינה רלוונטית, כי ההכרתיות-משמעותיות של הגירויים התחושתיים אופפת את היחיד כל הזמן, ללא הרף, כשעצם הימצאותו במצב של הכרתיות, של היות-חי, היא המשמעות הגדולה של החיים. כאמור להכרתיות ולמשמעות דרגות שונות וכתוצאה מתהליך האדפטציה עשויה משמעות I-מולדת לרדת בעוצמתה. במקרים כאלה מחפשים בני האדם שינוי מסוים המשנה את הגירויים התחושתיים או מעצים אותם; לכן בני אנוש יוצאים לנופשים, לטיולים, לבתי הקפה, למסיבות, הולכים לקולנוע ותיאטרון, צורכים מופעי תרבות וכיוצא בזה.

בנוסף לכך מפתחים בני האדם משמעויות II-רגילות ומשמעויות II-קיצוניות. משמעויות אלו מוקנות לאדם על ידי החברה הן בצורת ידע כללי ומדעי (יש חשיבות גדולה לכך שהיחיד יודע איך להתנהג בחברה שלה הוא שייך) והן בצורת אמונה שלה מתחייב היחיד באופן קוגניטיבי-אמוציונאלי. פיתוח של משמעויות אלו מתחיל כשהאדם הינו תינוק וממשיך לאורך כל ימי חייו, כאשר בגילים מתקדמים יותר (למשל, מגיל ההתבגרות ואילך) מתחיל היחיד להגיב באופן עצמאי למשמעויות II רגילה ולמשמעויות II-קיצוניות (אמונה דתית, חילוניות, אידיאולוגיה פוליטית וכו).

כללו של דבר, לפי מודל ה CM היחיד מוקף בכל שניה בהכרתיות ומשמעויות שונות בין שהן מולדות, נרכשות-רגילות, או בין שהן נרכשות-קיצוניות. אותו הסבר יינתן גם לשאלת האבסורד: רוב רובם של בני האדם אינם מוצאים את החיים אבסורדיים משום שחייהם מעוגנים במשמעויות חיים לסוגיהן ולדרגותיהן השונות. הסבר זה ניתן להכללה גם לתצפית האומרת שרוב רובם של בני האדם אינם מבצעים התאבדויות. מדוע? כי לחייהם יש משמעויות רבות. (וכאן איני רואה צורך לדון בתופעות של דכדוך ואף של דיכאון קל שתוקפים בני אדם מדי פעם מסיבות שונות, היות ומצבים אלה חולפים ובני האדם חוזרים למצב של תפקוד נורמאלי.) הבעיה שנשארה לפני המודל הנוכחי היא להסביר את המקרים המועטים שבהם היחיד מאבד את משמעות-החיים, תופס את העולם כאבסורד ומנסה להתאבד.

אתחיל בשאלה הבסיסית: כיצד מתמודד מודל ה CM עם איבוד משמעות-החיים? איך קורה הדבר שאדם מאבד משמעות I-מולדת, משמעות II-רגילה, ומשמעות II-קיצונית? אתחיל עם משמעות I-מולדת. אם משמעות זו היא מולדת, קשה להעלות על הדעת מקרים שבהם משמעות זו נעלמת, כלומר, קשה לחשוב על מצבים שבהם האדם שרוי בהכרתיות אולם המשמעות של קליטת הגירויים התחושתיים אובדת, נעלמת – כיצד זה ייתכן? אני סבור שמשמעות I-מולדת עלולה להיעלם במצבים של משברים קשים ביותר. מתוך התבוננות בחיים אני יכול לומר שישנם מצבים כאלה והם קשורים, כאמור, במחלות קשות מלאות כאבים וייסורים, מחלות חשוכת מרפא. במקרים אלה משמעות החיים הופכת לשלילית ואדם מבקש את נפשו למות – פשוט אין בכוחו לסבול עוד. אני מכיר מקרים של אנשים זקנים מאוד שחלו בסרטן פראי שגרם להם ייסורים קשים ואיבוד שליטה על תפקודי הגוף וכתוצאה יחלו למותם. אנשים אלה השלימו עם המוות, אמרו בליבם שאת החיים שלהם הם כבר חיו והגיע הזמן לשים קץ לכאב המתמשך והחסר כל תקווה. מקרים אחרים, כפי שכבר רמזתי לעיל, קשורים במות קרוב משפחה יקר ביותר. בישראל ידועים כמה מקרים שבהם אבות התאבדו על קבר הבן שנפל חלל במערכות ישראל. בכל המקרים האלה אפשר לאפיין את המצב בצורה הבאה: המשמעות השלילית גוברת על המשמעות החיובית. איך זה קורה מבחינה פסיכולוגית-קוגניטיבית? הינה ההסבר שאני מציע.

במצב קשה זה, עוסקת המערכת הקוגניטיבית רק במצבים מנטליים הקשורים בכאב ובייסורים, מצבים שמשמעותם הינה שלילית. תחילה מגייסת המערכת את כל משאביה להפחית את הכאב והייסורים, כשגיוס זה גורם להפחתה ולצמצום של קליטת גירויים תחושתיים, שבדרך כלל גורמים למשמעות חיובית (למשל, בני אדם ממעטים לצאת מחדרם, ממעטים לאכול, אינם קוראים ספרות, סולדים מראיית טלוויזיה ומשמיעת רדיו, נוטים להתבודד, מעדיפים אפלולית על פני אור וכו). לאחר זמן מה, עוסקת המערכת אך ורק בגירויים ובמחשבות השליליות כשייצוגו של העולם הרחב הולך ונעלם ומצטמצם לכדי ההתמקדות בכאב ובייסורים המתמשכים. אני סבור שבמקרים קוגניטיביים-מנטאליים כאלה מאבד היחיד במהירות את טעם החיים והוא מתחיל לראות בהמשך קיומו אבסורד – מה הטעם בהמשך המצב הנורא הזה שבו אני מצוי? והרהורים על התאבדות כמוצא מהמצב הקשה להחריד מתחילים לפעפע בראשו של האחוז בייסורים.

מה קורה בשעת משבר קשה למשמעות II (רגילה, קיצונית)? סביר להניח שאם במצבים הקשים האלה עוצמתה של משמעות I-מולדת יורדת ועלולה להיחלש, בוודאי שגם עוצמת משמעות II תרד ואף עלולה להיעלם. אני משער שבמצבי משבר קשים כאשר ההכרתיות מעמעמת את האצלת המשמעות על הגירויים התחושתיים (משום שהמערכת הקוגניטיבית מעבדת את האינפורמציה באופן מינימאלי) יאבדו מכוחן גם משמעויות II משום שהמערכת הקוגניטיבית פשוט תתעלם מהן. למשל, קשה להעלות על הדעת שאדם הסובל מסרטן לבלב המייסר אותו בכאבים נוראיים ימשיך להתעניין במימוש של תכנית כלכלה סוציאלית או ברפורמה במערכת החינוך, תחומים שהיו מרכז עיסוקו הקודם של אותו אדם. אף איש לא יתפלא אם החולה יהיה ממוקד כל כולו בשאלה הקריטית מתי הוא יקבל כבר זריקת מורפיום נוספת. אולם היחסים בין שני סוגי המשמעויות I ו- II אינם כה פשוטים. אני סבור שישנם לפחות שלושה מצבים בהם משמעות II עשויה לגבור על משמעות I.

1. *משמעות I כמטרה למשמעות II*: במקרה הנוכחי בני אדם הופכים את החשיפה לגירויים תחושתיים, את משמעויות I-מולדות, למטרות חייהם, לדרך חיים שבה ילכו, למסע של רדיפה אחר עינוגי החושים, כלומר, להעצמה מקסימלית של משמעות I-מולדת. צורת חיים זו הופכת מהר מאוד לחיי הוללות הקשורה בהימורים, חיי מין סוערים, שתיה וצריכת סמים. כל דרך חיים אחרת שהחברה מעוניינת לטפח אותה, נדחית בקלות רבה, כי משמעות I כדרך חיים מתממשת בקלות רבה על ידי עינוגי החושים. כפי שמציעה האמרה העממית: 'אכול ושגול כי מחר נפול'. אולם דרך זו מביאה במהירות רבה לחורבנו (הפיזי והנפשי) של היחיד ועליו לעשות משהו דרסטי ביותר כדי לעצור, להפסיק את מחול השדים שכל החיים מתמצים בעינוגי החושים, ולמצוא דרך חיים חדשה, משמעות חיים חדשה, משמעות II-קיצונית, שבעזרתה עשוי היחיד, העומד על סף התהום, להיוושע. ואכן בישראל ידועים מספר מקרים שבהם אנשי בוהמה התדרדרו לדרגה כזו שרק המעבר לחיי דת קיצוניים (יהדות חרדית) הצילו אותם מאבדון.
2. *בסוף החיים מתחזקת האמונה הדתית*: עם הזיקנה עוצמת הגירויים התחושתיים הולכת ונחלשת ומתעמעמת. למשל, נוצרת ירידה בכושר המיני, ביכולת הראייה והשמיעה, עולה לאות רבה וירידה בחשיפה לגירויים תחושתיים וגם באלה המופשטים. משמעות I-מולדת הולכת ומתערפלת וכתוצאה מחפש היחיד מפלט במשמעות II-נרכשת כשהדת מציעה את עצמה במהירות ובקלות גדולה והאדם עשוי להיכנע לה ולראות בה דרך חיים ולהפכה אף למשמעות II-קיצונית.
3. *ויתור על משמעות I בגלל מחויבות למשמעות II-קיצונית*: מדובר במקרים של מאבקים אידאולוגיים, פוליטיים וצבאיים הדורשים הקרבתו של היחיד למען מימוש מטרותיהם, כגון, מטרות לאומיות בזמן מלחמה, והפגנות אלימות נגד השלטון הרודני המושחת. במקרים מסוג זה עשוי היחיד לעיתים להקריב את עצמו למען המטרה, כלומר, היחיד מוותר על משמעות I למען משמעות II-קיצונית. למשל, בשעת מלחמה נעשים מעשי גבורה שבהם מקריב היחיד את עצמו למען המולדת, וכדי להציל את חבריו ממות וודאי. אלה הינם מקרים יוצאי דופן שהחברה ורוב בני האדם בארצות רבות משבחים ומציינים אותם כמעשי גבורה. אולם ישנם מספר מקרים שבהם אנשים מקריבים את חייהם למען אידיאלים מסוימים (הידועים לשמצה על ידי רוב בני האדם) כמו למשל תופעת הקמיקזה של היפנים נגד הצבא האמריקאי בימי מלחמת העולם השנייה, או תופעת השאהידים המוסלמים נגד הכופרים בני השטן – האמריקנים (השטן הגדול) והישראלים (השטן הקטן). (איני עומד לדון בסיבות הפסיכולוגיות-חברתיות העומדות ביסוד מעשים אלה ולעמוד על ההבדלים הגדולים בין הסוגים השונים של תופעות חברתיות אלו. אין זה המקום להרחיב בעניין זה. כל מה שאני מעוניין לעשות הוא להביא דוגמה ברורה שבה משמעות II-קיצונית גוברת על משמעות I-מולדת.)

אולם, כפי שנאמר לעיל, ברוב המקרים כשאדם עובר משבר קשה מתמוטטת משמעות II (רגילה, קיצונית) אולם משמעות I-מולדת אינה מתחסלת והיא ממשיכה לתמוך באדם כדי שישרוד את המשבר וימשיך בחייו. כדוגמה קלאסית לכך אדון עתה במשבר הקשה שעבר בזמנו הסופר הרוסי הגדול טולסטוי.

1. **הווידוי והנובלה של טולסטוי**

בשנת 1869 פורסם הרומן הענק של טולסטוי *מלחמה ושלום;* וב 1877 פורסם הרומן הגדול *אנה קארנינה.* טולסטוי היה אז בשיא תהילתו הספרותית וגם חיי משפחתו וניהול האחוזה שלו, יסנאיה פוליאנה, התנהלו היטב! ואז ב 1882 הוא ניסה לפרסם את ה*ווידוי* שלו, שבו הוא מדווח על כך שמשמעות-החיים אבדה לו וכל מה שהשיג עד כה בחייו נראה לו חסר ערך לחלוטין, אם כך, עבור מה לחיות? (הכנסייה האורתודוקסית אסרה את פרסום הספר שיצא לאור בשלמותו בז'נבה ב 1884 וברוסיה ב 1906.) כעבור כמה שנים, ב 1886, פורסמה הנובלה *מותו של איוון איליץ'* המתאר את גיבור הסיפור, איוון, כאדם שמשמעות חייו אבדה לו וכתוצאה הוא עובר ייסורים קשים ביותר.

מה הביא את טולסטוי למצב אומלל זה, למשבר אמצע החיים? (ראו דיון מעניין בנושא זה אצל לוריא, 2002.) מתברר שהמשבר נעוץ במחשבה שאין לחיים משמעות משום שאלוהים אינו קיים והמוות הקרב הוא הדבר היחידי הוודאי. מחשבה זו באה לידי ביטוי חריף ב*ווידוי* של טולסטוי שמתאר בדפיו הראשונים את המשל הבא המרמז על חיי טולסטוי עצמו. וזהו המשל שאכנהו 'הבאר': אדם בורח על נפשו ונופל לתוך באר. בתחתית הבאר פוער את פיו דרקון אכזרי. תוך כדי נפילה תופס האדם שיח הצומח מהקיר ובכך מציל את חייו. אבל, אבוי, שני עכברים אחד לבן ואחד שחור מכרסמים את גזע השיח והאדם מבין שקיצו קרב כי השיח יקרע והוא ייפול למלתעות הדרקון. לפתע מבחין האדם שעל פרחי השיח מופיעות טיפות דבש והוא מלקק אותם. כאמור טולסטוי מייחס את המשל לעצמו. הוא נאחז בשיח, כלומר בחיים, אחוז בפחד מפני המוות הוודאי, דהיינו, הדרקון. הוא אינו מבין מדוע הוא נקלע למצב נורא זה ומנסה להתנחם בלקיקת הדבש, אבל הטעם אינו מהנה אותו והעכברים, היום והלילה, מכרסמים את גזע השיח שבו נאחז טולסטוי (משך החיים המוגבל הולך ואוזל). ובכן, מה קרה לטולסטוי? לפי המשל הוא נקלע למצב שבו הנאות החיים (הדבש) פגו כי הדבר הבטוח הוא המוות ההולך ומתקרב (העכברים מכרסמים את השיח ובסופו של דבר ייפול האדם, טולסטוי, אל מותו, לפיו של הדרקון) ואין טולסטוי יכול להתנחם באמונה באלוהים, כי הוא איבד אותה מזמן. (כאן אני מוכרח להתוודות ולומר שכשקראתי את סיפור 'הבאר' חשבתי שהאדם יסלק את העכברים וישליך אותם לפיו של הדרקון, אבל לא כך התרחשו הדברים במשל זה.) ובכן, אם החיים אינם אלא בחזקת ליקוק דבש עד המוות הוודאי, איך יש להתייחס למצב כזה? איך על האדם להגיב נוכח המוות הבטוח ללא נחמה שסיפקה עד כה האמונה באל?

וכך כותב טולסטוי (2015) ב*ווידוי* שלו: "השאלה שלי – זאת שבגיל חמישים הביאה אותי אל סף ההתאבדות – הייתה השאלה הפשוטה ביותר הטמונה בנשמתו של כל אדם, החל בילד חסר דעת וכלה בזקן החכם ביותר – אותה שאלה שבלעדיה לא ייתכנו חיים. כפי שנוכחתי לדעת על בשרי. עיקר השאלה הוא: 'מה ייצא ממה שאני עושה כעת, ממה שאעשה מחר – מה ייצא מכל חיי?' בניסוח אחר, השאלה תיראה כך: 'לשם מה עלי לחיות, לשם מה עלי לרצות משהו, לשם מה עלי לעשות משהו?' דרך נוספת לבטא את השאלה היא: 'האם יש בחיי משמעות, שהמוות המצפה לי באופן בלתי נמנע לא ימחה אותה?'" (עמ' 48).

טולסטוי חיפש תשובות לשאלתו הבסיסית על משמעות-החיים ולא הצליח לקבל תשובה מספקת לא על ידי הדת ולא על ידי המדע. יתר על כן, הוא מצא שני הוגים גדולים שהגיעו כמוהו לאותה שאלה על משמעות-החיים, שופנהאואר, הפילוסוף הגרמני וקהלת, שלמה המלך, שאותו הוא מצטט בהרחבה ('הבל הבלים הכל הבל' וכן הלאה.)

כאן מן הראוי להדגיש שהמשבר הנוכחי לא נחת על טולסטוי לפתע פתאום, אלא הוא נבנה על יסודות של חייו בעבר. מתברר שטולסטוי סבל מייסורי מצפון קשים ומטרידים שנים רבות מגיל הבחרות. הוא גינה את חייו שהיו הוללים, פרועים, פרוצים ובלתי מוסריים בעליל וקיווה לגאולה בנשואים ובהקמת משפחה. יתר עלכן, בגיל השמונה עשרה איבד לדבריו את האמונה באלוהים. עובדה אחרונה זו העלתה בי את התמיהה הבאה: אם טולסטוי הפסיק להאמין באלוהים בגיל הבחרות, מדוע התעורר בו משבר משמעות-החיים שמעוגן באיבוד האמונה לאחר כל כך הרבה זמן, כעבור כשלושים שנה, בסביבות גיל החמישים? האם יתכן שעל אף חוסר האמונה באלוהים מבחינה תבונית, שכנו בליבו של טולסטוי מספר יסודות של מוסר נוצרי, משום שאחרת קשה יהיה להבין את ייסורי המצפון של הסופר הדגול ואת רצונו להיות אדם טוב?

טולסטוי מציע ארבע דרכים אפשריות לטיפול בבעיית משמעות-החיים:

1. *אי-מודעות למוות*: אם אדם אינו מודע למותו הוודאי, המוות לא יטריד אותו. אולם, טוען טולסטוי ברגע שהאדם מודע לכך אין אפשרות להתעלם מידיעה מחרידה זו. (ושוב אני מוכרח להעלות שאלת תם, האם טולסטוי, שהשתתף במלחמת קרים בצעירותו, נהיה מודע למוות רק בגיל חמישים?)
2. *הנאה*: כשאדם מודע לכך שהחיים הם קצרים וסופיים, יבקש ליהנות מהם ככל האפשר. טולסטוי דוחה אפשרות זו בגלל שהיא נראית לו חסרת מוסר. (ושוב תמיהה, הרי טולסטוי עד לזמן שהתחתן בגיל השלושים המוקדמות, היה ידוע כהולל גדול.)
3. *התאבדות*: זו דרך שטולסטוי מצא בה ביטוי ליושר אינטלקטואלי, אבל מציין בכנות רבה שהוא חסר אומץ לבצע אותה.
4. *המשך החיים*: למרות שמשמעות-החיים אבדה לו, מציין טולסטוי שהוא ממשיך לחיות הלאה באופן בלתי נסבל ואבסורדי.

בסופו של דבר יוצא טולסטוי ממשבר זה כאיש מאמין, אולם אמונתו אינה עולה בקנה אחד עם האמונה הנוצרית פרבוסלבית על שלל טקסיה שרווחו בימי חייו (טולסטוי דחה טקסים אלה ומספר טיעונים ורעיונות אחרים באמונה זו). הוא הטיף (ממש כך) לחיים מעשיים פשוטים בנוסח של המוז'יקים (האיכרים) ברוסיה משום שרק כך מתבצע רצון האל והחיים מקבלים משמעות. חיי המעשה הפשוטים הם שמביאים לדעתו לגאולה הדתית, לאמונה באל שבאה מתוך הקרביים, מהאמוציות, מעצם חיי המעשה והעמל כמו הפשוטים שבפשוטי העם. לדעתו רק כך נעשה רצונו של אלוהים, רק מחיי עמל תמימים, עולה האמונה באל, אמונה שבלעדיה אין משמעות-חיים. וכך כתב טולסטוי ב*ווידוי* שלו: "אם האדם חי, הרי יש בהכרח משהו שהוא מאמין בו. אלמלא האמין שיש בשביל מה לחיות , לא היה חי. (עמ' 66). והוא המשיך וכתב, "... אך ניצלתי רק בזכות זאת, שהספקתי לפרוץ את גבולות הייחודיות שלי, לחזות בחייו של עם העמלים הפשוט ולהבין שאלה הם החיים האמתיים היחידים." (עמ' 74). והוסיף ואמר, "הרי אינני חי כשאני מאבד אמונה בקיומו של אלוהים, הרי כבר מזמן הייתי שולח יד בנפשי אלמלא הייתה לי תקווה עמומה למצוא אותו. הרי אני חי, חי באמת ובתמים. ... לדעת את אלוהים ולחיות – זה הינו הך. אלוהים הנו חיים." (עמ' 76). שינוי זה בדרך החיים הביא בהכרח לקונפליקטים קשים לא רק עם הכנסייה הפרובוסלבית, אלא גם עם משפחתו ובמיוחד עם אשתו, קונפליקט שקרע את משפחתו בריב בלתי מתפשר. בסופו של דבר, בזקנתו ברח טולסטוי מביתו, התקרר מאוד בנסיעה ברכבת ומת שם בתחנת רכבת (ב 1910).

אפשר להעלות מספר קווי דמיון מעניינים בין (א) קאמי שכתב את הרומן *הזר* ולאחר מכן את המסה הפילוסופית *המיתוס של סיזיפוס* שמסבירה את משבר משמעות-החיים והאבסורד שבחיים שבאו לידי ביטוי ברומן שלו *הזר*, ושמצא לבסוף גאולה בסולידאריות האנושית שבאה לידי ביטוי *ברומן הדבר* ובמסה *הפילוסופית האדם המורד,* לבין (ב) טולסטוי שכתב תחילה את *הווידוי* המתאר ומסביר את משבר משמעות-החיים והאבסורד שבחיים של טולסטוי עצמו ולאחר מכן את הרומן *מותו של איוון איליץ'* המביא לידי ביטוי ספרותי את משבר-החיים של טולסטוי עצמו. איוון איליץ' היה שופט מצליח, אגואיסט שביקש לעצמו אך חיים קלים ונוחים ושהתייחס לבני משפחתו כאל מטרד נסבל ללא כל אמפתיה. הוא חלה לפתע במחלה מוזרה שהחלה בכאב קל בצד הגוף והתפתחה למחלה שהביאה לידי סופו של איוון.לנוכח המוות המתקרבאל גיבור הרומן, עובר איוון משבר משמעות-חיים שבו כל חייו הקודמים נראים לו כחסרי כל ערך, ללא כל משמעות (בדומה לטולסטוי ב*ווידוי*). בסופו של דבר מגיעה הגאולה של איוון: ברגעים האחרונים של חייו, כשהוא מתייחס באופן אנושי ומבין את נקודות הראות של בנו הצעיר החומל עליו מעומק ליבו, הוא מרגיש אמפתיה וחמלה לבנו וגם לאשתו. ואז נעלמת חרדתו הגדולה מפני המוות והוא מרגיש שמחה שלאחריה הוא נפטר בשלווה. בדומה לקאמי שמצא גאולה באחווה האנושית (ה*דבר* ו*האדם המורד*) מצא איוון איליץ' (כלומר טולסטוי) שלווה כתוצאה מתפיסת משמעותם הגדולה של בני משפחתו שהיו חסרי כל חשיבות עבורו קודם לכן.

בהיתן תיאורים קצרים אלה של שתי יצירותיו של טולסטוי, ה*ווידוי* ו*מותו של איוון איליץ'*, אני חוזר להשערתי הקודמת: ברוב המקרים כשאדם עובר משבר קשה מתמוטטת משמעות II (רגילה, קיצונית) אולם משמעות I-מולדת אינה מתחסלת וממשיכה לתמוך בהמשך חייו. מה שתומך בהשערה זו היא עצם העובדה הפשוטה שטולסטוי עצמו מודה בה שהוא המשיך בחייו על אף האבסורד שבחיים אלה. וכך כותב טולסטוי ב*ווידוי*: "כל הנימוקים האלה לא יכלו לשכנע אותי עד כדי כך, שאעשה את מה שנבע מטיעוניי, כלומר, אשלח יד בנפשי. הייתי אומר דבר שקר, אילו אמרתי שהגעתי למה שהגעתי בעזרת התבונה, ולא הייתי שולח יד בנפשי. התבונה פעלה את פעולתה, אך פעל גם משהו נוסף, שאיני יכול לכנותו אלא *תודעת החיים*." (עמ' 62. ההדגשה היא שלי ובאה להבליט את קו הדמיון בין 'תודעת החיים' לבין 'משמעות I-מולדת'.) יתר על כן, טולסטוי עצמו מצא לעצמו משמעות II-קיצונית חדשה שהביאה לו דרגה מסוימת של שלווה פנימית – הוא הפך לאדם דתי בעל צבע מעין נזירי, הוא חי חיים פשוטים כמו מוז'יק רוסי. בגלל דרך חיים חדשה זו שלו הוא שאף לחלק את רכושו בין פשוטי העם, דבר שכאמור הביא לידי קרע נורא בינו לבין אשתו סופיה אנדרייבנה ולמריבות קשות במשפחתו.

ובכן איוון, טולסטוי וגם קאמי נאחזו בחיים עד לרגע האחרון (כמו האדם במשל ה'באר') ומצאו נחמה באחווה האנושית, כשטולסטוי עצמו מצא דרך חיים חדשה, מפלט בדת הנוצרית הפשוטה של המוני העם. מבחינה זו יש לציין שקאמי אמנם מצא גאולה בסולידאריות האנושית, אולם ללא כל נימה דתית שאותה הוא לא קיבל. למשל, בעוד שהאב פָּנָלוּ ברומן *הדֶבֶר טוען* שהדֶבֶר בא על מאזיניו כעונש מצד האל ומציע להם לחזור אל מתחת לכנפי הכנסייה, אל חיקה, מוחה הרופא רייֶה נגדו ואומר לו שהדֶבֶר המית ילד חף מכל חטא, ומעלה את השאלה על מה נענש הילד, איזה חטא הוא חטא? לגבי רייֶה הדת אינה אלא התכחשות לאמת והוא אינו מקבל אותה בשום פנים ואופן. לבסוף מת פנלו ממחלה שלא זוהתה.

המשבר של טולסטוי מעלה מספר תמיהות שעל חלקן רמזתי לעיל. ראשית, טולסטוי הפסיק להאמין באלוהים בגיל צעיר למדי, בגיל שמונה-עשרה. אם כך, מדוע אי אמונה זו החלה להטריד אותו דווקא בגיל כה מאוחר, בגיל החמישים? מה קרה בגיל זה, מהו אותו גורם מסוים שאליו הצטרפה אי האמונה והביאה את טולסטוי לידי משבר משמעות-החיים?

שנית, טולסטוי יצא ממשבר חייו במין התגלות של אמונה נוצרית ודרך חיים המבוססת על חייהם הפשוטים של המוז'יקים באחוזתו. ושוב עולה שאלת הזמן: טולסטוי חי במשך מרבית חייו בסביבתם של מוז'יקים והיה עד לחייהם הקשים והמרודים. אם כך, מדוע עמד טולסטוי על כך פשטות חיים אלה לפתע פתאום רק בגיל החמישים?

שלישית, אם פשוטי העם, הבורים שאינם יודעים קרוא וכתוב, מבינים את האמת על אודות האלוהים, כיצד ייתכן שטולסטוי האיש האינטליגנטי והמשכיל לא הבין דבר כזה פשוט וברור? האם ההשכלה והיהירות של טולסטוי סמאה את עיניו והוא התפכח רק בגיל חמישים?

רביעית, אדם משכיל ואינטליגנטי כמו טולסטוי היה צריך להבין שדורות קודמים של פשוטי העם ברוסיה ובעולם האמינו בכל מיני אמונות שונות לחלוטין מהנצרות הפרובוסלבית (ומהנצרות בכללה על כל גווניה השונים); האם כל פשוטי העם האלה לדורותיהם הבינו את האמת האלוהית שעליה דיבר טולסטוי? התשובה השלילית ברורה, ולכן ברור שאמונתם של פשוטי העם אינה מעידה דבר על האמת האלוהית. אם כך, מה גרם לטולסטוי להתעלם משיקול דעת זה? (וכאן אני חייב לציין שה*ווידוי* שזור בוויכוחים מוסריים-דתיים מסוג זה שהציקו מאוד לטולסטוי.)

את התשובה לשאלות אלו ניתן למצוא במאמר שכתבתי ב 2016 על הרומן של טולסטוי: *אנה קארנינה*. במאמר זה העליתי את ההשערה הבאה על אודות משבר משמעות-החיים של טולסטוי. (מצאתי להשערה זו מספר תמיכות בספרות המקצועית הדנה בטולסטוי, אך לא ב*ווידוי* – שם לא מצאתי התייחסות לכתיבת שני רומנים אלה הנחשבים עד היום לפסגת הפרוזה.) לאחר כתיבתם של שני הרומנים הריאליסטיים הגדולים שלו (*מלחמה ושלום, אנה קארנינה*) מיצה טולסטוי את כל מקורות הכתיבה שלו המבוססים על חייו הוא.מה עוד יכול היה טולסטוי לכתוב שלא כתב בהרחבה בשני רומנים אלה? כדי לכתוב רומן ריאליסטי רחב יריעה חייב הסופר להישען על מאגר עצום של זיכרונות שממנו יוכל לדלות בקלות רבה חומרי גלם כדי לעצב גיבורים ועלילה. (כך גם נכתבו הרומנים הנפלאים של צ'רלס דיקנס ושל ג'ק לונדון, רומנים שניזונו מאוסף עצום של זיכרונות העבר של יוצריהם.) אולם לאחר כתיבת שני הרומנים האלה התייבשו המקורות של טולסטוי וכוח כתיבתו הריאליסטי נדם. הוא היה חייב למצוא מקור חדש שממנו ישאב חיים ליצירתו החדשה. כתוצאה נכנס טולסטוי למשבר יצירתי שערער אותו לחלוטין. למשבר זה הצטרפו שני הגורמים הנ"ל, אי האמונה הדתית וחייהם הקשים של המוז'יקים (שאליהם הצטרף כנראה גם רגש אשם של הרוזן טולסטוי שחי על חשבון עמלם). הפתרון שטולסטוי מצא למשבר חייו היא בחזקת דרך חיים המנוגדת לזו שהוא ניהל עד למשבר: הוא הפך להיות דתי, ניהל אורך חיים כמעט נזירי כשל איכר רוסי, דחה את יצירתו הספרותית הקודמת והחל לכתוב ספרות המגויסת לרעיונותיו החדשים.