פרק 1.1 מתוך הספר "אופי התודעה" מאת דייוויד צ'אלמרס,

הוצאת אוניברסיטת אוקספורד, 2010

1

התמודדות עם בעיית התודעה

**1. מבוא**

התודעה מעוררת את הבעיות הסבוכות ביותר במדע של המיינד [mind]. אין דבר שאנו מכירים כה מקרוב כמו חווית התודעה, אך גם אין דבר שקשה יותר להסביר. כל סוגי התופעות המנטליות נחקרו בשנים האחרונות באופן מדעי, אך בפני חקר התודעה ניצבו קשיים רבים. רבים ניסו להסביר את התודעה, אך ההסברים מחטיאים תמיד את המטרה. עניין זה הוביל אחדים להניח שבעיית התודעה אינה פתירה ושלא ניתן לתת לה שום הסבר שמניח את הדעת.

על מנת להתקדם עם בעיית התודעה, עלינו להתעמת עמה ישירות. בפרק זה אני מבודד תחילה את החלק הקשה באמת של השאלה, על ידי כך שאני מפרידו מהחלקים הפתירים יותר ומנמק מדוע כל כך קשה להסבירו. אני מבקר מספר מחקרים חדשים שמשתמשים בשיטות רדוקטיביות כדי לטפל בנושא התודעה וטוען שאין זה מן הנמנע ששיטות אלה אינן מצליחות להתגבר על החלק הקשה של הבעיה. ברגע שמזהים את הכישלון הזה נפתח צוהר להתקדמות נוספת. במחצית השנייה של הפרק אני טוען שאם נמצא סוג חדש של הסבר לא-רדוקטיבי ניתן יהיה לתת הסבר נטורליסטי לתודעה.

**2. הבעיות הקלות והבעיה הקשה**

אין רק בעיה אחת של התודעה. "תודעה" היא מושג רב-משמעי שמתייחס לתופעות רבות ושונות. יש להסביר את כל אחת מהתופעות הללו, אולם חלק מהן ניתן להסביר בקלות רבה יותר. בהתחלה יהא זה יעיל לחלק את הבעיות הקשורות לתודעה ל"קשות" ו"קלות". הבעיות הקלות של התודעה הנן אלה שנדמה כי ניתן להחיל עליהן ישירות את השיטות הסטנדרטיות של המדעים הקוגניטיביים, שיטות אשר עמן מסבירים תופעה במונחים של מנגנונים חישוביים או עצביים. הבעיות הקשות הן אלה שנדמה כי לא ניתן להחיל עליהן את השיטות הללו.

הבעיות הקלות של התודעה נוגעות להסבר התופעות הבאות:[[1]](#footnote-1)

* היכולת להבחין בגירויים סביבתיים, לסווגם לקטגוריות ולהגיב להם
* שילוב מידע באמצעות מערכת קוגניטיבית
* היכולת לדווח על מצבים מנטליים
* יכולתה של מערכת לגשת למצביה הפנימיים
* מיקוד תשומת הלב
* השליטה המכוּונת בהתנהגות
* ההבדל בין ערנות ושינה

כל התופעות הללו מזוהות עם מושג התודעה. לעיתים אנחנו אומרים שמצב מנטלי הוא מודע כשניתן לדווח עליו באופן מילולי או כשהוא נגיש מבפנים. לעיתים אנחנו אומרים שמערכת היא מודעת למידע מסוים כאשר יש לה את היכולת להגיב על בסיס המידע הזה, או באופן ברור יותר, כשהיא משיגה מידע זה או כשהיא מסוגלת לשלבו ולנצלו תוך שליטה מתוחכמת בהתנהגות. אנחנו אומרים לעיתים שפעולה היא מודעת בדיוק כשהיא נעשית בכוונה תחילה. לעיתים קרובות אנחנו אומרים שאורגניזם מסוים נמצא בהכרה, כלומר הוא מודע, כדי לומר במילים אחרות שהוא ער.

אין שום ספק לגבי השאלה אם ניתן להסביר את התופעות *הללו* באופן מדעי. את כולן ניתן להסביר ללא קושי במונחים של מנגנונים חישוביים או עצביים. על מנת להסביר את הנגישות לתודעה ואת יכולת הדיווח על התודעה עלינו לזהות רק את המנגנון שבאמצעותו נוצר מידע על מצבים פנימיים ונעשה אפשרי לדיווח לשוני. כדי להסביר את שילוב המידע עלינו רק להציג מנגנונים שבאמצעותם נאסף מידע המנוצל לתהליכים מאוחרים יותר. כדי להסביר שינה וערנות יספיק הסבר נוירופיזיולוגי ראוי של התהליך האחראי להתנהגות האורגניזם במצבים המנוגדים הללו. בכל מקרה, ברור שדגם קוגניטיבי או נוירופיזיולוגי מתאים יכול לספק את ההסבר.

לו התופעות הללו היו כל מה שיש לתודעה להציע, אז התודעה לא הייתה מהווה ממש בעיה. על אף שעדיין אין לנו שום דבר שיהא קרוב להסבר מלא של התופעות הללו, יש לנו רעיון ברור לגבי איך ניגשים להסבר שלהן. משום כך אני מכנה אותן הבעיות הקלות. כמובן ש"קל" הוא מושג יחסי. כדי להגיע לפרטים מדויקים יידרשו ככל הנראה מאה או מאתיים שנה של עבודה אמפירית קשה. אך קיימות סיבות טובות להאמין שהמתודות של המדעים הקוגניטיביים ומדעי המוח ינחלו הצלחה.

הבעיה הקשה באמת של התודעה היא בעיית ההתנסות. כשאנו חושבים משהו או תופסים אותו באופן חושי יש מין קול שקשוק של עיבוד מידע, אך יש גם היבט סובייקטיבי. כפי שנייגל [Nagel] (1974) ניסח זאת, יש "משהו שהוא כמו להיות" *[*something it is like to be*]* אורגניזם מודע. ההיבט הסובייקטיבי הזה הוא התנסות. כשאנו רואים אנו *מתנסים* בתחושות ויזואליות: האיכות המורגשת של ההיות-אדום, ההתנסות של חושך ואור, האיכות של עומק בשדה ויזואלי. התפיסה החושית במודוסים השונים מחוללת התנסויות אחרות: צליל הקלרינט, הריח של כדורי העש. באורגניזם קיימות גם תחושות גופניות כתוצאה מכאב; דימויים מנטליים פנימיים שהועלו באוב; האיכות המורגשת של האמוציות; והחוויה של זרם חשיבה מודעת. מה שמאחד את כל המצבים הללו הוא שיש "משהו שהוא כמו להיות" בהם. הם כולם מצבים של התנסות.

ללא ספק, אורגניזמים מסוימים הם סובייקטים של התנסות, אולם אין זה ברור דיו מהי הסיבה לכך שמערכות אלו הן סובייקטים של התנסות. מדוע כאשר המערכות הקוגניטיביות שלנו עוסקות בעיבוד מידע חזותי ושמיעתי, יש לנו חוויה חזותית שמיעתית: האיכות של כחול עמוק, התחושה של צליל דוֹ? כיצד נוכל להסביר מדוע "יש משהו שהוא כמו" להחזיק בדימוי מנטלי או לחוות רגש? הרוב מסכימים שהתנסות נוצרת על ידי יסוד פיזיקלי, אך אין לנו שום הסבר טוב מדוע ואיך היא נוצרת כך. מדוע עיבוד תהליכים פיזיקליים מחולל חיים פנימיים עשירים? מבחינה אובייקטיבית נדמה שהוא לא אמור לחולל זאת, ובכל זאת הוא מחולל זאת.

אם קיימת איזושהי בעיה שניתן לכנותה בעייתה המובהקת של התודעה, הרי שזו הבעיה הזאת. על פי המשמעות המרכזית של מושג ה"תודעה", אורגניזם הוא מודע אם קיים "משהו שהוא כמו להיות" האורגניזם הזה, ומצב מנטלי הוא מודע אם קיים "משהו שהוא כמו להיות" במצב הזה. כדי לציין זאת משתמשים לעיתים גם במונחים כמו "תודעה פנומנלית" ו"קוואליה", אך אני חושב שזה טבעי יותר לדבר על "התנסות מודעת" או פשוט "התנסות". יש עוד דרך יעילה להימנע מבלבול (הדרך שבה משתמשים Newell 1990; Chalmers 1996) והיא לשמור את מונח ה"תודעה" עבור התודעה של ההתנסות, ולהשתמש במונח הטעון פחות "מודעוּת" עבור התופעות הפשוטות יותר שתוארו לעיל. אם כולם היו מחליטים להסכים לכך באופן נרחב, התקשורת הייתה הרבה יותר קלה; אולם כפי שהדברים נראים כרגע, אלה שמדברים על "תודעה" מדברים לרוב על דברים שונים לגמרי.

פילוסופים ומדענים שכותבים על הנושא מנצלים בדרך כלל את חוסר הבהירות לגבי מושג ה"תודעה". קל למצוא מאמרים בנושא תודעה שפותחים בהכרזה חגיגית בדבר המסתורין של התודעה, ובה הם טורחים לציין את מוזרוּת הסובייקטיביות כביכול, הואל ולא ניתן לתפסה או לבטאה, ומודאגים מכך שעד כה אין ברשותנו תיאוריה של התופעה. הנושא אצלם הוא כמובן הבעיה הקשה – בעיית ההתנסות. בהמשך המאמר אנו שומעים קולות אופטימיים יותר והמחברים מציגים את התיאוריה שלהם בעניין התודעה. אחרי בחינת התיאוריה הזו היא מתגלה כתיאוריה של אחת מהתופעות הפשוטות ביותר – של היכולת לדווח, כלומר של גישה אינטרוספקטיבית, או איך שלא יכנו זאת. בשלב הסיום המחברים מצהירים שהסתבר שהתודעה התגמשה בסופו של דבר, אולם הקוראים נותרים בתחושה שהם נפלו קורבן לטקטיקת המכירה המניפולטיבית שבה מפתים את הלקוח לקנות מוצר מסוים ובסוף מוכרים לו משהו אחר. הבעיה הקשה נותרת אצל המחברים האלו ללא מענה.

**3. הסבר תפקודי**

מדוע הבעיות הקלות הנן קלות, והבעיה הקשה - קשה? הבעיות הקלות הנן קלות בדיוק מפני שהן נוגעות להסבר של יכולות ותפקודים קוגניטיביים. כל שאנו זקוקים לו כדי להסביר תפקוד קוגניטיבי הוא לזהות מנגנון שיכול לבצע את התפקוד הזה. השיטות של המדעים הקוגניטיביים מתאימות היטב לסוג זה של הסבר ולכן הן מתאימות היטב לבעיות הקלות של התודעה. בניגוד לכך, הבעיה הקשה היא קשה בדיוק מפני שהיא אינה בעיה של ביצוע תפקודים. הבעיה נותרת על כנה גם כשמסבירים את ביצועם של כל התפקודים הרלבנטיים. (לא נעשה כאן שימוש במילה "תפקוד" במובן הטלאולוגי הצר של משהו שמערכת נועדה לעשות, אלא במובן הרחב של כל תפקיד שמערכת מסוימת מבצעת ושגורם לה לייצר התנהגות מסוימת.)

להסביר את היכולת לדווח למשל, פירושו אינו אלא להסביר כיצד מערכת יכולה לבצע את התפקיד של ייצור דיווחים על מצבים מנטליים. כדי להסביר גישה פנימית עלינו להסביר כיצד מערכת יכולה להיות מושפעת על ידי מצביה הפנימיים ולהשתמש במידע לשם וויסות תהליכים מאוחרים יותר. כדי להסביר את השילוב של מידע ואת השליטה בו עלינו להסביר כיצד תהליכים מרכזיים במערכת מסוגלים לאסוף מידע ולהשתמש בו כדי לפשט פעולות שונות. כל אלה הן בעיות הנוגעות להסבר של תפקודים.

כיצד אנו מסבירים ביצוע של תפקוד? על ידי זיהוי של מנגנון שמבצע את התפקוד.[[2]](#footnote-2) התכנון של מודלים נוירופיזיולוגיים וקוגניטיביים הוא מושלם עבור המשימה הזו. אם ברצוננו למצוא הסבר מפורט, אך ברמה נמוכה, באפשרותנו לזהות את המנגנון העצבי שאחראי לתפקוד. אם אנחנו מעוניינים בהסבר יותר מופשט, אנחנו יכולים לזהות מנגנון במונחים חישוביים. כך או כך, התוצאה תהיה הסבר מלא ומשביע-רצון. ברגע שזיהינו את המנגנון העצבי או החישובי שמבצע למשל את התפקוד של דיווח לשוני, עיקר עבודתנו בנוגע להסבר של היכולת לדווח הסתיים.

במובן מסוים, הנקודה הזו מובנת מאליה. לגבי התופעות הללו, זו עובדה מושגית שההסבר שלהן כולל רק הסבר של תפקודים שונים, הואיל והתופעות ניתנות להגדרה מבחינה תפקודית. העובדה שהיכולת לדיווח מגולמת במערכת מסוימת פירושה אינו אלא שלמערכת יש את היכולת לדיווח לשוני על מידע פנימי. העובדה שמערכת מסוימת נמצאת במצב ערני פירושה אינו אלא שהמערכת מסוגלת לקבל מידע מהסביבה ולהשתמש בו לצורך הנחיית ההתנהגות בדרך הנכונה. על מנת להיווכח שמדובר בעובדה מושגית יש לשים לב לכך שמישהו שאומר "הסברת את הביצוע של תפקוד דיווח לשוני, אך לא הסברת את היכולת לדיווח" מבצע טעות מושגית טריוויאלית בנוגע ליכולת דיווח. כל מה שאפשר לדרוש לשם הסבר של היכולת לדיווח הוא הסבר כיצד התפקוד האמור מבוצע; הדבר נכון גם לגבי התופעות האחרות שהוזכרו.

ההסבר הרדוקטיבי פועל בכל המדעים הגבוהים בדיוק בצורה הזאת. כדי להסביר למשל את הגֶנים היינו צריכים לזהות את המנגנון ששומר ומעביר מידע תורשתי מדור לדור. מה שהתגלה זה שה-DNA מבצע את התפקוד הזה; ברגע שאנחנו מסבירים כיצד התפקוד מתבצע הסברנו את הגנים. כדי להסביר מה זה חיים אנחנו בהכרח חייבים להסביר כיצד מערכת יכולה להתרבות, להסתגל לסביבתה, לחלף חומרים וכיוצא באלה. כל אלה הן שאלות על ביצוע תפקודים ולכן מתאימות היטב להסבר רדוקטיבי. הדבר נכון לגבי רוב הבעיות במדעים הקוגניטיביים. כדי להסביר מהי למידה עלינו להסביר את הדרך שבה יכולות התנהגותיות של המערכת משתנות לאור מידע סביבתי ואת הדרך שבה מידע חדש יכול להשפיע על התאמת פעולות המערכת לסביבתה. ברגע שאנחנו יודעים כיצד מנגנון עצבי או חישובי מסוים מבצע את העניינים הללו, הסברנו למידה. אנחנו יכולים לומר את אותו הדבר על תופעות קוגניטיביות אחרות כגון תפיסה, זיכרון, ושפה. לפעמים צריך לתאר את התפקודים הרלבנטיים באופן די מורכב, אבל מה שברור זה שברגע שהמדעים הקוגניטיביים מבקשים להסביר את התופעות האלו, הם עושים זאת על ידי הסבר של ביצוע התפקודים.

כשמדובר בהתנסות מודעת, סוג כזה של הסבר חסר. מה שהופך את הבעיה הקשה לקשה וכמעט ייחודית זה העובדה שהיא הולכת מעבר לבעיות הנוגעות לביצוע תפקודים. כדי להבין זאת יש לתת את הדעת על כך שגם אם הסברנו את ביצועם של כל התפקודים הקוגניטיביים וההתנהגותיים הקשורים להתנסות – הבחנה תפיסתית, קטגוריזציה, גישה אל הפְנים, דיווח לשוני – נותרת שאלה נוספת ללא מענה: מדוע ביצועם של תפקודים אלו מלוּוה בהתנסות? גם אם מסבירים את התפקודים השאלה הזו נותרת פתוחה.

אין בנמצא עוד שאלה כזו כאשר מסבירים מהם גנים או חיים או למידה. אם מישהו אומר "אני מבין שהסברת כיצד DNA שומר ומעביר מידע תורשתי מדור לדור, אך לא הסברת למה זה גן", אז הוא מבצע טעות מושגית. הפירוש של להיות גן אינו אלא להיות ישות שמבצעת את התפקיד של השימור וההעברה האמורים. אולם אם מישהו אומר "אני מבין שהסברת כיצד מנתחים מידע, כיצד משלבים אותו ומדווחים עליו, אך לא הסברת כיצד מתנסים בכך", הוא אינו מבצע טעות מושגית. זוהי שאלה נוספת שאינה מובנת מאליה.

השאלה הנוספת הזו היא שאלת מפתח בבעיית התודעה. מדוע עיבודו של כל המידע הזה אינו מתרחש "באפלה", כלומר בהעדר כל רגש פנימי? מדוע כאשר צורות גל אלקטרומגנטיות נטבעות ברשתית העין ומובחנות וממוינות על ידי מערכת חזותית, ההבחנה והמיון נחוות כתחושה של אדום חי? אנו יודעים שהתנסות מודעת אכן נוצרת כשתפקודים אלה מתבצעים, אולם עצם העובדה שהיא נוצרת היא התעלומה המרכזית. קיים פער הסברי [explanatory gap] (מושג שטבע לוין [Levine], 1983) בין התפקודים וההתנסות, ואנו זקוקים לגשר הסברי כדי להתגבר עליו. כל עוד אנחנו מסבירים באופן סתמי את התפקודים בלבד, אנחנו נשארים רק בצד אחד של הפער ההסברי ועל כן עלינו לחפש את החומרים עבור הגישור עצמו דווקא במקום אחר.

אין זה אומר שההתנסות עצמה היא נטולת תפקודים. אולי יסתבר שיש לה תפקודים חשובים, אך בהתעלם מהתפקודים שעשויים להיות לה, ההסבר שלה מורכב יותר מההסבר הפשוט של תפקודים. אולי אף יסתבר שבמהלך ההסבר של התפקודים עצמם, נגיע לתובנה הבסיסית שתאפשר להסביר מהי התנסות. אולם גם אם זה יקרה, התגלית שנקבל תהיה גמול נוסף בצורת הסבר. אין בנמצא תפקוד קוגניטיבי כזה שנוכל לומר מראש שההסבר לתפקוד הזה יסביר לנו אוטומטית מהי התנסות.

אנו זקוקים לגישה חדשה כדי להסביר התנסות. השיטות הרגילות של הסברים במדעים הקוגניטיביים ובמדעי המוח אינן מספיקות. שיטות אלה פותחו כדי להסביר ביצוע תפקודים קוגניטיביים בלבד והן עובדות מצוין לצורך כך. עם זאת, כפי ששיטות אלה עובדות כיום הן מסוגלות להסביר רק את ביצועם של תפקודים. כשעוסקים בבעיה הקשה הגישה המקובלת אינה עוזרת במאומה.

**4. מקרי בוחן אחדים**

בשנים האחרונות עסקו מחקרים אחדים במדעים הקוגניטיביים ובמדע המוח בבעיות התודעה. עובדה זו עשויה להוכיח שהניתוח שלפנינו שגוי, אולם למעשה בדיקה מדוקדקת של העבודות האמורות רק מעניקה תמיכה נוספת לניתוח שלי. בבואנו לחקור את היבטי התודעה שמחקרים אלה בודקים ואילו היבטים הם מסבירים בסופו של דבר, אנו מגלים שהמטרה הסופית של ההסברים היא תמיד אחת מהבעיות הקלות. אני מדגים זאת בעזרת שתי דוגמאות מייצגות.

הראשונה היא "תיאוריה נוירו-ביולוגית של התודעה" שפותחה על ידי קריק וקוך (Crick and Koch 1990; ראו גם Crick 1994). תיאוריה זו מתמקדת בתנודות עצביות מסוימות בתדירות 75-35 הרץ באזור קליפת המוח; ההיפותזה של קריק וקוך הייתה שתנודות אלו הן בסיס התודעה. אחת הסיבות לכך היא שלתנודות אלו יש קורלציה עם המודעוּת במספר מודוסים שונים – למשל במערכות הראייה והריח – וגם מפני שנדמה שהן מהוות מנגנון שבעזרתו ניתן לאחות [bind] מידע. איחוי הוא התהליך שבו פיסות מידע נפרדות אודות ישות אחת מצורפות יחדיו כדי לשמש בעיבודים מאוחרים יותר, כמו כשמאחדים מידע על צבע וצורה של עצם מסוים שנקלט בחושים מתוך ערוצים ויזואליים שונים. ההיפותזה של קריק וקוך, שהולכים בעקבות חוקרים אחרים (למשל Eckhorn et al. 1988), היא שאיחוי מידע נעשה על ידי סינכרוניזציית תנודותיהן של קבוצות העצבים המייצגות את התכנים הרלבנטיים. כאשר שתי פיסות מידע אמורות להיכרך יחדיו הקבוצות העצביות הרלבנטיות יתנודדו באותה תדירות ופאזה.

הפרטים לגבי איחוי המידע וכיצד הוא נעשה עדיין אינם מובנים כראוי, אך הבה נניח שאפשר להגיע להבנה לגביהם. מה עשויה להסביר התיאוריה שתתקבל? מן הסתם היא תסביר את איחוי המידע, ואולי אף תספק הסבר כללי יותר של שילוב מידע במוח. קריק וקוך גם טוענים שתנודות אלו כנראה מפעילות את המנגנונים של זיכרון העבודה, כך שאולי יהיה בקרוב הסבר עבורו ואולי גם עבור צורות זיכרון אחרות. בסופו של דבר, התיאוריה אולי תוביל להסבר כללי של קישור מידע שנקלט בחושים ונשמר בזיכרון לצורך שימוש בעיבוד מאוחר יותר.

תיאוריה כזו הנה בעלת ערך, אך היא לא תאמר לנו דבר אודות הסיבה להתנסות בתכנים המדוברים. קריק וקוך טוענים שהתנודות האלו הן המִתאמים [correlates] העצביים של ההתנסות. טענה זו נתונה במחלוקת – האם איחוי אינו קורה גם בעיבודו של מידע לא מודע? – אך גם אם מסכימים עליה, השאלה ההסברית נותרת על כנה: מדוע התנודות יוצרות התנסות? הבסיס היחיד לקשר הסברי הוא התפקיד שיש להן באיחוי ושמירת מידע, אבל לעולם אין התייחסות לשאלה מדוע התהליכים עצמם של איחוי ושמירה מלוּוים בהתנסות? אם איננו יודעים מדוע איחוי ושמירה יוצרים התנסות, הסיפור על התנודות אינו יכול לסייע לנו. לעומת זאת, אם היינו יודעים מדוע איחוי ושמירה יוצרים התנסות, אז הפרטים הנוירופיזיולוגיים היו רק עיטור הקצפת שעל העוגה. התיאוריה של קריק וקוך חבה את קיומה להנחה שקיים קשר בין איחוי והתנסות ולכן היא אינה יכולה להסביר בעצמה את הקשר הזה.

אינני סבור שקריק וקוך טוענים שהם אכן מצאו מענה לבעיה הקשה של התודעה, על אף שיש פרשנויות אחדות בכיוון זה. ראיון שפורסם עם קוך מכיל הצהרה ברורה בדבר המגבלות של שאיפות התיאוריה:

הבה נשכח קודם את ההיבטים הקשים באמת, כמו רגשות סובייקטיביים, הואיל וכנראה שאין להם פתרון מדעי. המצב הסובייקטיבי של משחק, כאב, הנאה, של לראות כחול, של להריח ורד – נדמה שיש פער גדול בין הרמה החומרית, ההסבר של מולקולות ונוירונים, והרמה הסובייקטיבית. הבה נתמקד בדברים שקל יותר ללמוד – כמו מודעות ויזואלית. עכשיו אתה מדבר איתי, אבל אתה לא מסתכל עליי, אתה מסתכל על הקפוצ'ינו וכך אתה מודע לקיומו. אתה יכול לומר "זהו ספל ויש בו נוזל". אם אתן לך אותו, אתה תזיז את ידך ותיקח אותו – לאופן תגובתך תהיה משמעות רבה. את זה אני מכנה בשם מודעות. (מהי תודעה? *Discover* [נובמבר 1992] 96).

הדוגמה השנייה היא גישה פסיכולוגית קוגניטיבית. זוהי תיאוריית חלל-העבודה [workspace] הגלובלית של התודעה על פי ברנארד בארס [Baars], אשר מוצגת בספרו תיאוריה קוגניטיבית של התודעה. על פי תיאוריה זו, תכני התודעה נכללים בחלל עבודה גלובלי, שהוא מעבד מרכזי הנועד לתווך תקשורת בין המוני מעבדים מומחים לא-מודעים. כאשר המעבדים המומחים האלה צריכים לשגר מידע לשאר המערכת, הם עושים זאת על ידי שליחת המידע לחלל העבודה שפועל כלוח מודעות קהילתי עבור שאר המערכת. לוח מודעות זה נגיש לכל יתר המעבדים.

בארס משתמש במודל זה כדי להתייחס להיבטים רבים של המצב האנושי וכדי להסביר מספר ניגודים בין תפקוד קוגניטיבי מודע ולא מודע. יחד עם זאת, מדובר בתיאוריה של נגישות קוגניטיבית שמסבירה כיצד תכני מידע מסוימים הנם הנגישים ביותר בתוך מערכת, כמו גם בתיאוריה של שילוב מידע ויכולת דיווח על מידע. נראה שכתיאוריה של המודעות, כלומר של הקורלציה התפקודית של התנסות מודעת, זו יכולה להיות תיאוריה מוצלחת, אולם היא אינה מציעה את ההסבר להתנסות עצמה.

ניתן בטעות לחשוב שלפי התיאוריה הזו תכני ההתנסות הנם בדיוק תכניו של חלל העבודה. אף על פי כן, גם אם זהו המצב, התיאוריה הזו אינה מסברירה כלל מדוע מתנסים במידע שנמצא בחלל העבודה הגלובלי. הדבר הטוב ביותר שהתיאוריה הזו יכולה לעשות הוא לומר שמתנסים במידע מפני שהוא נגיש באופן גלובלי. אולם כעת השאלה עולה בניסוח אחר: מדוע נגישות גלובלית מחוללת התנסות מודעת? כמו תמיד, שאלה זו, אשר אמורה להוות גשר לנושא ההתנסות, נותרת ללא מענה.

כמעט על כל המחקרים בשנים האחרונות שנוקטים בגישה קוגניטיבית לתודעה, או בגישה של מדעי המוח, ניתן למתוח ביקורת דומה. מודל ה"דרווניזם העצבי" של אדלמן (1989), למשל, עוסק בשאלות הנוגעות למודעות תפיסתית ולתפיסה העצמי, אך אינו אומר דבר על הסיבה שבגינה אמורה להיווצר התנסות. מודל "הסקיצות הרבות" של דנט (1991) שואף בעיקר להסביר את היכולת לדווח על תכנים מנטליים מסוימים. תיאוריית ה"רמה הבינונית" של ג'אקנדוף (1988) מספקת הסבר לכמה תהליכים חישוביים שעומדים ביסוד התודעה, אבל ג'אקנדוף מדגיש שהשאלה אודות האופן שבו הם "משליכים" על התנסות מודעת נותרת עלומה.

חוקרים המשתמשים בשיטות אלה לרוב אינם מתבטאים באופן מפורש בנוגע לגישות שלהם כלפי בעיית ההתנסות המודעת, על אף שלעיתים הם נוקטים בעמדה ברורה. אפילו בקרב אלה שנוקטים בעמדה ברורה כלפי בעיה זו הגישות שונות מאוד. כשבוחנים את סוג העבודות האלה בהקשר של בעיית ההתנסות, מגלים שקיימות אסטרטגיות שונות. זה היה יעיל יותר אם הבחירות האסטרטגיות האלה היו נעשות בצורה מפורשת יותר.

האסטרטגיה הראשונה היא פשוט להסביר משהו אחר. חוקרים אחדים טוענים במפורש שכרגע בעיית ההתנסות היא מסובכת מדי ואולי אף נמצאת מחוץ לתחום המדע. במקום זאת, חוקרים אלה מעדיפים להתייחס לאחת הבעיות הקלילות יותר כמו יכולת הדיווח או מושג העצמי. על אף שכיניתי את הבעיות הללו "קלות" הן שייכות לבעיות הבלתי-פתירות והמעניינות ביותר במדעים הקוגניטיביים, כך שהעבודות הללו הנן בעלות ערך רב. הדבר הגרוע ביותר שניתן לומר על הבחירה שלהן הוא שבמסגרת חקר התודעה זו בחירה יחסית נטולת שאיפות, ולעיתים ניתן לפרש לא נכון את העבודות הללו.

הבחירה השנייה היא לנקוט בגישה מחמירה יותר ולהכחיש את התופעה (גרסאות שונות של גישה זו הוצעו על ידי Allport 1988; Dennett 1991; Wilkes 1988). לפי גישה זו, ברגע שהסברנו את התפקודים כמו נגישות ויכולת דיווח, לא קיימת עוד תופעה הנקראת "התנסות" שיש להסבירה. אחדים מכחישים באופן מפורש את התופעה בתואנה שמה שלא ניתן לאמת באופן חיצוני אינו יכול להיות מציאותי. אחרים מגיעים לאותה מסקנה על ידי כך שהם טוענים שהתנסות קיימת, אולם רק אם אנחנו מזהים את ה"התנסות" כמעין יכולת ליצור הבחנות ולדווח על מצבים מנטליים. גישות אלה אמנם מובילות לתיאוריה פשוטה, אך הן לגמרי לא משביעות רצון. התנסות היא ההיבט המרכזי והבולט ביותר של החיים המנטליים שלנו והיא כנראה אולי מה שהכי דורש הסבר [key explanandum] במדע של המיינד. עקב הסטטוס הזה שלה, בבחינת מה שמספק הסברים, לא ניתן לשלול את קיום ההתנסות כשיוצרים תיאוריה חדשה בתואנה שהיא הרוח החיה. אדרבה, היא העובדה המרכזית שכל תיאוריית תודעה מוכרחה להסביר. תיאוריה שמכחישה את התופעה "פותרת" את הבעיה על ידי התחמקות מהשאלה.

האפשרות השלישית של החוקרים היא לטעון שהם מסבירים את ההתנסות במובנה המלא. חוקרים אלה (בניגוד לאלה שהוזכרו לעיל) מבקשים לקחת את ההתנסות ברצינות רבה; הם משרטטים את המודל או התיאוריה התפקודיים שלהם וטוענים שהם מסבירים את האיכות הסובייקטיבית המלאה של ההתנסות (למשל Flohr 1992; Humphrey 1992). אולם הדבר האופייני להם הוא שהם במהרה פוסחים על השלב הרלבנטי ובדרך כלל הוא נראה בסוף כמו קסם. לאחר שמספקים כמה פרטים אודות עיבוד מידע, ההתנסות נכנסת לפתע לתמונה, אך זה לא ברור כיצד תהליכים אלה מחוללים התנסות. כנראה שפשוט לוקחים כמובן מאליו את העובדה שהם מחוללים זאת, אבל אז ההסבר אינו מלא וזו מעין גרסה של האסטרטגיה החמישית שתוזכר בהמשך.

הגישה הרביעית, המוצלחת יותר, נוקטת בשיטות אלה כדי להסביר את מבנה ההתנסות. אפשר לטעון למשל שההסבר עבור ההבחנות, שהמערכת הוויזואלית יוצרת, יכול להסביר את הקשרים המבניים בין התנסויות צבע שונות, וגם את המבנים הגאומטריים של השדה הוויזואלי (ראו למשל Clarck 1992; Hardin 1992). באופן כללי, עובדות מסוימות בנוגע למבנים, אשר נמצאו בעיבוד המידע, יתאימו וכנראה יסבירו עובדות הנוגעות למבנה ההתנסות. אסטרטגיה זו מתקבלת על הדעת, אך היא מוגבלת. היא לוקחת, לכל היותר, את קיומה של ההתנסות כמובן מאליו ומסבירה מספר עובדות על המבנה שלה, תוך שהיא מספקת הסבר לא-רדוקטיבי של היבטיה המבניים של ההתנסות (אדבר על כך יותר בהמשך). אמנם זה יעיל למטרות רבות, אך אין זה אומר דבר על הסיבה שבגינה אמורה בכלל להיות התנסות.

האסטרטגיה החמישית והנבונה היא לבודד את מצע ההתנסות. ככלות הכול, כמעט כל אחד טוען שתהליכים במוח מחוללים התנסות בדרך זו או אחרת, וזה הגיוני לזהות את סוג התהליך שמחולל אותה. דוגמאות לכך מהווים קריק וקוך, אשר מציגים את עבודתם כמבודדת את המתאם העצבי של התודעה, ואדלמן (1989) וג'אקנדוף (1987), אשר טוענים טענות דומות. ביסוס הטענות הללו דורש ניתוח תיאורטי זהיר, במיוחד משום שהתנסות אינה ניתנת להתבוננות ישירה בניסויים, אך אם משתמשים באסטרטגיה זו כראוי היא יכולה לשפוך אור בעקיפין על בעיית ההתנסות. אף על פי כן, האסטרטגיה הזו אינה מושלמת. כדי להגיע לתיאוריה משביעת רצון אנחנו צריכים לדעת יותר מאשר רק לדעת איזה תהליך מחולל התנסות; אנו זקוקים גם להסבר מדוע וכיצד. תיאוריה מלאה של התודעה חייבת לבנות גשר הסברי.

**5. המרכיב הנוסף**

ראינו שיש סיבות מתודולוגיות מדוע השיטות של המדעים הקוגניטיביים ומדעי המוח אינן מצליחות להסביר התנסות מודעת. אלה הן פשוט הסוג הלא נכון של מתודות. הן אינן יכולות להניב כלל הסבר מתאים. כדי להסביר התנסות מודעת, אנו זקוקים למרכיב נוסף בהסבר. עניין זה מאתגר את אלה שרציניים לגבי השאלה הקשה של התודעה: מהו המרכיב הנוסף שלך, ומדוע הוא אמור להסביר התנסות מודעת?

קיים שפע של מרכיבים נוספים שאפשר לגבב איכשהו. יש חוקרים שמציעים מרפא לבעיה בצורת שילוב בין כאוס ודינמיקה לא-לינארית. יש שחושבים שהמפתח טמון בעיבוד לא-אלגוריתמי. יש שמאמינים בתגליות עתידיות בנוירופיזיולוגיה. אחרים מניחים שהמפתח לתעלומה נמצא ברמה של מכניקת הקוונטים. קל לראות מדוע הוצעו כל הרעיונות האלו. אף אחת מהשיטות הישנות אינן עובדות, ולכן הפתרון חייב להימצא במשהו חדש. למרבה הצער, בכל ההצעות הללו התגלעו אותן הבעיות הישנות.

פנרוז [Penrose] (1989, 1994) הציע למשל עיבוד לא-אלגוריתמי על סמך התפקיד שכנראה יש לו בתהליכים של הבנה מתמטית מודעת. הטיעונים בנושא מתמטיקה שנויים במחלוקת, אולם גם אם הם ינחלו הצלחה ויהיה הסבר לעיבוד לא-אלגוריתמי במוח האנושי, עדיין יהא זה הסבר הנוגע רק לתפקודים הדרושים לחשיבה מתמטית, וכיוצא באלה. עבור עיבוד לא-אלגוריתמי כמו גם עבור עיבוד אלגוריתמי, השאלה שנותרת ללא תשובה היא: מדוע עיבוד זה מחולל התנסות? בתשובה לשאלה זו אין לעיבוד הלא-אלגוריתמי תפקיד מיוחד.

אותו הדבר נכון גם לדינמיקה לא-לינארית וכאוטית. היא עשויה לספק הסבר חדשני לדינמיקה של תפקוד קוגניטיבי, קצת שונה מזה הניתן על ידי השיטות המקובלות במדעים הקוגניטיביים. אולם מדינמיקה אחת מקבלים אך ורק עוד דינמיקה. גם בגישה זו, השאלה אודות ההתנסות היא מסתורית כהרגלה. הנקודה הזו אף ברורה יותר בהקשר של תגליות חדשות בנוירופיזיולוגיה. תגליות אלו עשויות לעזור לנו להתקדם ממש בהבנת תפקודי המוח, אבל אותה השאלה תתעורר ביחס לכל תהליך עצבי שאנו מבודדים. קשה לדמיין איזה חלק בנוירופיזיולוגיה חדשה צפוי להתגלות ולחרוג מעבר להסבר של תפקודים קוגניטיביים חדשים. זה הרי לא שנגלה פתאום איזה זוהר פנומנלי בתוכו של איזה נוירון!

ה"מרכיב הנוסף" הפופולרי ביותר מכולם הוא כנראה מכניקת הקוונטים (למשל, Hameroff 1994). האטרקטיביות של תיאוריית הקוונטים של התודעה נובעת אולי מתוך החוק הנוגע למינימיזציה של תעלומות: התודעה היא תעלומה, ומכניקת הקוונטים היא תעלומה, כך שאולי לשתי התעלומות יש מקור משותף. אף על פי כן, בתיאוריות הקוונטים של התודעה מתגלים אותם קשיים כמו בתיאוריות עצביות או חישוביות. לתופעות הקוונטיות יש בהחלט כמה תכונות תפקודיות מרשימות, כמו חוסר-דטרמיניזם והעדר-מרחביוּת. טבעי לדמיין שלתכונות האלו יש תפקיד כלשהו בהסבר של תפקודים קוגניטיביים מסוימים כמו בחירה אקראית ושילוב מידע, ולכן היפותזה זו לא יכולה להישלל על הסף באופן אפריורי. אך כשמדובר בהסבר עבור ההתנסות, התהליכים הקוונטיים נמצאים באותה הסירה, ממש כמו כל היתר. השאלה מדוע תהליכים אלה אמורים לחולל התנסות נותרת לחלוטין ללא מענה.[[3]](#footnote-3)

(המשיכה המיוחדת לתיאוריות קוונטים נובעת מכך שלפי פרשנים אחדים של מכניקת הקוונטים לתודעה יש תפקיד פעיל ב"מיטוט" פונקציית גל הקוונטים. פרשנויות מעין אלה נתונות במחלוקת, אבל בכל מקרה הן אינן פותחות פתח להסבר של התודעה במונחים של תהליכי קוונטים. אדרבה, תיאוריות אלה מניחות את קיומה של התודעה ומשתמשות בהנחה זו במסגרת ההסבר של תהליכים קוונטים. לכל היותר, תיאוריות אלה אומרות לנו משהו על התפקיד הפיזיקלי שיש לתודעה. אולם הן אינן אומרות לנו דבר על הסיבה שבגינה היא נוצרת).

בסופו של דבר, אותה הביקורת חלה על כל הסבר פיזיקלי טהור של התודעה. לכל תהליך פיזיקלי שנזהה תהיה שאלה שלא נענתה: מדוע אמור תהליך זה ליצור התנסות? כל תהליך כזה יכול להתרחש גם בהעדר התנסות, והעובדה הזו היא קוהרנטית מבחינה מושגית. מכאן נובע שאף הסבר של התהליך הפיזיקלי הטהור לא יגיד לנו מדוע התנסות נוצרת. עובדת היווצרות ההתנסות חורגת מעבר למה שניתן להסיק מתיאוריה פיזיקלית.

הסבר פיזיקלי טהור פירושו הסבר של מבנים פיזיקליים, כלומר הסבר של מבנים מקרוסקופיים במונחים של יחידות מיקרו-מבניות שונות. הסבר כזה מספק גם הסבר ראוי לביצועם של תפקודים, כלומר הסבר עבור התפקודים האלו במונחים של מנגנונים פיזיקליים, אשר מבצעים אותם. זה קורה משום שברגע שהנתונים הפנימיים של הסבר פיזיקלי נתונים, התכונות המבניות והתפקודיות מתגלות כתוצאה אוטומטית מהנתונים. עם זאת, המבנה והדינמיקה של תהליכים פיזיקליים מניבים רק עוד מבנים ודינמיקות, לכן אנחנו יכולים לצפות לכך שמבנים ותפקודים הם הדבר היחיד שהתהליכים האלה יסבירו. לעומת זאת, העובדות בנוגע להתנסות אינן יכולות להיות תוצאה אוטומטית של איזשהו הסבר פיזיקלי, הואיל והעבדה שכל תהליך נתון יכול להתקיים גם ללא התנסות היא קוהרנטית מבחינה מושגית. התנסות עשויה להיווצר מתוך עניין פיזיקלי, אך היא אינה מוסברת על ידי העניין הפיזיקלי.

מה שניתן ללמוד מכך זה שלא ניתן להסביר התנסות מודעת בעבור זוטות. זו עובדה מעניינת ששיטות רדוקטיביות – שיטות שמסבירות תופעות מורכבות אך ורק במונחים של תהליכים פיזיקליים בסיסיים – נוחלות הצלחה בתחומים רבים. ניתן להסביר את רוב התופעות הביולוגיות והקוגניטיביות בעבור פרוטות, אם מבינים את התופעות כתוצאות אוטומטיות של תהליכים בסיסיים יותר. זה היה יכול להיות נהדר אם שיטות רדוקטיביות היו יכולות להסביר גם התנסות; במשך זמן רב קיוויתי שהן תוכלנה. למרבה הצער, יש סיבות מתודולוגיות לכישלונן המתחייב של שיטות אלו. שיטות רדוקטיביות מצליחות ברוב התחומים מפני שמה שיש להסביר בתחומים אלו זה מבנים ותפקודים, וזהו סוג הדברים שהסבר פיזיקלי יכול לכלול. כשזה מגיע לבעיה שחורגת מעבר להסבר של מבנים ותפקודים שיטות אלו נעשות עקרות.

נדמה כי טענתי מהדהדת את הטענה הויטליסטית ששום הסבר פיזיקלי לא יכול להסביר חיים, אולם אלה לא מקרים דומים. מה שהניע את הספקנות הויטליסטית היה ספק לגבי יכולתם של מנגנונים פיזיקליים לבצע את התפקודים המדהימים הרבים שקשורים לחיים, כמו התנהגות סתגלנית מורכבת והתרבות. הטענה המושגית, לפיה מה שנדרש הוא הסבר עבור התפקודים, התקבלה באופן לא מפורש, אולם הואיל ולא היה בידי הויטליסטים ידע רב אודות מנגנונים ביוכימיים, הם הטילו ספק באפשרות שתהליכים פיזיקליים יכולים לבצע את האמור והציעו את ההיפותזה של הרוח החיה כהסבר אלטרנטיבי.[[4]](#footnote-4) ברגע שהתגלה שתהליכים פיזיקליים מסוגלים לבצע את התפקודים הדרושים נעלמו ספקותיהם של הויטליסטים כלא היו.

אולם כשמדובר בהתנסות, הסבר פיזיקלי של התפקודים אינו מה שנחוץ. להפך, המפתח טמון מושג ההתנסות עצמו שלפיו ההסבר של התפקודים אינו מספיק עבור ההסבר של ההתנסות באופן עקרוני. הנקודה המושגית הבסיסית אינה דבר שחקירות נוירו-מדעיות נוספות ישפיעו עליו. באופן אנלוגי, התנסות אינה דומה ל-èlan vital. הרוח החיה הוצגה כעיקרון הסברי על מנת להסביר את התפקודים הרלבנטיים ולפיכך ניתן לשלול אותה כאשר התפקודים האלה יוסברו בלעדיה. התנסות היא אינה עיקרון הסברי אלא מה שבעצמו ראוי להסבר ולכן היא אינה מועמדת לפסילה מסוג זה.

יש לציין שסוגים שונים של תופעות לא ברורות התגלו בסופו של דבר כמה שניתן להסבירו במונחים פיזיקליים. אך הן כולן היו בעיות הנוגעות להתנהגות הגלויה של עצמים פיזיקליים והסתכמו בבעיות בהסבר של מבנים ותפקודים. משום כך, תופעות אלה תמיד היו סוג הדברים שהסבר פיזיקלי עשוי להסביר, גם אם בכמה נקודות היו סיבות טובות לחשוד שאף הסבר כזה לא יגיע עוד בהמשך. היסק האינדוקציה המפתה מתוך מקרים אלו נכשל במקרה של תודעה, הואיל והיא אינה בעיה של מבנים פיזיקליים ותפקודים. בעיית התודעה מתמיהה באופן שונה לגמרי. ניתוח הבעיה יוכיח לנו שהסבר רדוקטיבי לחלוטין פשוט אינו יכול להסביר בהצלחה התנסות מודעת.

**6. הסבר לא-רדוקטיבי**

בנקודה זו יש שמתפתים לוותר בטענה שלעולם לא תהיה לנו תיאוריה של התנסות מודעת. מקגין [McGinn] (1989), למשל, טוען שהבעיה קשה מדי עבור שכלנו המוגבל; אנחנו "סגורים מבחינה קוגניטיבית" ביחס לתופעה. אחרים טענו שהתנסות מודעת כלל אינה בהישג ידה של תיאוריה מדעית.

לדעתי הפסימיות הזו מגיעה מוקדם מדי. זה לא המקום לוותר; זה המקום שבו העניינים מתחילים להיות מעניינים. כאשר המתודות הפשוטות של ההסברים נדחות עלינו לחקור את האלטרנטיבות. אם הסבר רדוקטיבי נכשל, הסבר *לא-רדוקטיבי* יהיה הבחירה הטבעית.

למרות שהתגלה שניתן להסביר לגמרי תופעות רבות מספור בעזרת ישויות שפשוטות יותר מהן, זה לא עניין אוניברסלי. בפיזיקה, יש לעיתים צורך לתפוס ישות מסוימת כפונדמנטלית. ישויות פונדמנטליות אינן מוסברות לעולם במונחים של משהו פשוט יותר. במקום זאת, תופסים אותן כבסיסיות ומפתחים תיאוריה המסבירה כיצד הן מתייחסות לכל דבר אחר בעולם. לדוגמה, במאה ה-19 התגלה שלא ניתן להסביר תהליך אלקטרומגנטי כתהליך מכני לחלוטין כפי שתיאוריות פיזיקליות קודמות טענו, לכן מקסוול ואחרים הציגו את המטען האלקטרומגנטי ואת הכוחות האלקטרומגנטיים כרכיבים פונדמנטליים חדשים של תיאוריה פיזיקלית. כדי להסביר אלקטרומגנטיות האונטולוגיה של הפיזיקה הייתה צריכה להתרחב. היה צורך בתכונות בסיסיות חדשות וחוקים בסיסיים כדי לתת הסבר מספק של התופעה.

מאפיינים נוספים שהתיאוריה הפיזיקלית תופסת כפונדמנטליים הם מסה וחלל-זמן. לא נעשה שום מאמץ להסביר את המאפיינים הללו על בסיס משהו יותר פשוט. ברם, זה לא שולל את האפשרות של תיאוריה של מסה או חלל-זמן. קיימת תיאוריה מורכבת המסבירה את האופן שבו מאפיינים אלו קשורים ביניהם ואת החוקים הבסיסיים שחלים עליהם. נוטים להשתמש בתיאוריה זו כדי להסביר תופעות מסובכות רבות הנוגעות למסה, חלל וזמן.

אני סבור שתיאוריית התודעה אמורה לתפוס את ההתנסות כפונדמנטלית. אנו יודעים שתיאוריית התודעה דורשת תוספת של משהו פונדמנטלי לאונטולוגיה שלנו, הואיל וכל דבר בתיאוריה פיזיקלית עולה בקנה אחד עם היעדר תודעה. היינו יכולים להוסיף מאפיין לא-פיזיקלי חדש לחלוטין שממנו ניתן להקיש על ההתנסות, אך קשה לדעת איזה מין מאפיין זה יהיה. סביר יותר שנבין את ההתנסות עצמה כמאפיין פונדמנטלי של העולם, לצד מסה, מטען, וחלל-זמן. אם נבין את ההתנסות כפונדמנטלית, אז אנחנו יכולים לגשת לעסוק בבניית תיאוריית תודעה.

היכן שיש תכונה פונדמנטלית ישנם חוקים פונדמנטליים. תיאוריה לא-רדוקטיבית של ההתנסות תוסיף עקרונות חדשים לחוקי הטבע החדשים. עקרונות יסוד אלו הם אלה שיישאו בנטל ההסבר במסגרת תיאוריית תודעה. בדיוק כשם שאנו מסבירים תופעות מורכבות כוללות, אשר מכילות מסה, על ידי עקרונות בסיסיים יותר, אשר מכילים מסה וישויות אחרות, ביכולתנו להסביר תופעות כוללות, אשר מכילות התנסות, על ידי עקרונות בסיסיים יותר, אשר מכילים התנסות וישויות אחרות.

מעל הכול, תיאוריה לא-רדוקטיבית של התנסות תגדיר עקרונות יסוד שאומרים לנו כיצד התנסות תלויה במאפיינים פיזיקליים של העולם. עקרונות פסיכופיזיקליים אלו לא יתערבו בחוקים הפיזיקליים, הואיל וחוקים פיזיקליים מהווים כבר מערכת סגורה. להפך, הם יהיו תוספת לתיאוריה פיזיקלית. תיאוריה פיזיקלית מספקת תיאוריה של תהליכים פיזיקליים, ותיאוריה פסיכופיזיקלית אומרת לנו כיצד תהליכים אלו מחוללים התנסות. אנחנו יודעים שהתנסות תלויה בתהליכים פיזיקליים, אך אנחנו גם יודעים שלא ניתן להסיק את התלות הזו אך ורק מחוקים פיזיקליים. עקרונות היסוד החדשים, שנוסחו על ידי תיאוריה לא-רדוקטיבית, מעניקים לנו את המרכיב הנוסף, שאנו מזדקקים לו כדי לבנות גשר הסברי.

אם תופסים את ההתנסות כפונדמנטלית, אז כמובן שהגישה הזו לא ממש אומרת לנו מדוע יש התנסות מלכתחילה, אבל זה נכון לגבי כל תיאוריה פונדמנטלית. הפיזיקה כלל אינה אומרת לנו מדוע יש חומר מלכתחילה, אולם איננו תופסים זאת כפגם של תיאוריות החומר. כל תיאוריה מדעית תופסת מאפיינים מסוימים של העולם כפונדמנטליים. תיאוריית חומר יכולה להסביר כל מיני עובדות על חומר, מעצם כך שהיא מוכיחה כיצד העובדות הללו הן מסקנות מחוקים בסיסיים. הדבר נכון גם לגבי תיאוריית ההתנסות.

עמדה זו מאופיינת כסוג של דואליזם, הואיל והיא קובעת תכונות בסיסיות שהולכות מעבר לתכונות שהפיזיקה מגלה. אלא שזו גרסה תמימה של דואליזם, אשר מתיישבת לגמרי עם השקפת העולם המדעית. גישה זו כלל אינה סותרת את התיאוריה הפיזיקלית; עלינו פשוט להוסיף עקרונות *מגשרים* נוספים כדי להסביר כיצד התנסויות נוצרות מתהליכים פיזיקליים. בתיאוריה הזו אין שום דבר רוחני או מיסטי – צורתה הכללית היא של תיאוריה פיזיקלית, עם ישויות פונדמנטליות אחדות המקושרות בחוקים פונדמנטליים. אמנם היא מרחיבה את האונטולוגיה במעט, אך גם מקסוול עשה זאת. המבנה הכללי של העמדה הזו הוא למעשה נטורליסטי לחלוטין, הואיל והיא טוענת שהיקום מסתכם ברשת ישויות בסיסיות המצייתות לחוקים פשוטים ושאולי יש תיאוריית תודעה המעוצבת על פי החוקים האלו. אם צריך לכנות עמדה זו בשם, אזי דואליזם נטורליסטי יהיה בחירה טובה.

אם השקפה זו נכונה, אז במובנים מסוימים תיאוריית תודעה דומה יותר לתיאוריה בפיזיקה מאשר לתיאוריה בביולוגיה. בתיאוריות ביולוגיות אין עקרונות שהם פונדמנטליים במובן הזה, ולכן יש בהן יותר אי-סדר ומורכבות מסוימת; אולם הואיל ותיאוריות בפיזיקה עוסקות בעקרונות פונדמנטליים הן שואפות לפשטות ואלגנטיות. חוקי הטבע הפונדמנטליים הנן חלק מהריהוט הבסיסי של העולם, ותיאוריות פיזיקליות אומרות לנו שהריהוט הבסיסי הזה הוא מדהים בפשטותו. אם תיאוריית תודעה כוללת גם עקרונות פונדמנטליים, אז עלינו לצפות לאותו דבר. עקרונות הפשטות, האלגנטיות, ואפילו היופי, אשר מניעים את חיפושם של הפיזיקאים אחר תיאוריה פונדמנטלית, חלים גם על תיאוריית התודעה.

(פילוסופים אחדים [מטריאליסטים מקבוצה ב' בפרק 5] טוענים שלמרות שיש פער מושגי בין תהליכים פיזיקליים והתנסות, אין פער מטאפיזי בהכרח, כך שהתנסות היא עדיין פיזיקלית במובן מסוים. העליתי טענות כנגד ההשקפה הזו בפרקים 7-5 ובפרק 10. עם זאת, אם מה שאמרתי עד כה נכון, עמדה זו מוכרחה לפחות להודות שיש פער הסברי בין תהליכים פיזיקליים והתנסות. למשל, לא ניתן יהיה להקיש על העקרונות הקושרים את הפיזיקלי וההתנסותי מתוך חוקי הפיזיקה, כך שיש לתפוס את העקרונות הללו כפונדמנטליים *מבחינת ההסבר*. לפיכך אפילו בהשקפה הזו, המבנה ההסברי של תיאוריית תודעה יהיה כפי שתיארתי.

**7. קווי מתאר של תיאוריית תודעה**

זה לא מוקדם מדי להתחיל לעבוד על תיאוריה. אנחנו כבר יכולים להבין עובדות מפתח אחדות אודות היחס בין תהליכים פיזיקליים והתנסות ואודות החוקיות שקושרת אותם. ברגע שמוותרים על הסבר רדוקטיבי, ביכולתנו להניח את העובדות האלה על השולחן כדי שהן תוכלנה לקבל את התפקיד הראוי להן בתור היחידות הראשוניות בתיאוריה לא-רדוקטיבית, ובתור רכיבים של חוקים בסיסיים המכוננים תיאוריה כוללת.

יש בעיה ברורה שמעיקה על התפתחות תיאוריית תודעה: דלות הנתונים האובייקטיביים. לא ניתן לחזות ישירות בהתנסות מודעת במסגרת ניסוי, ולכן איננו יכולים לאסוף נתונים אודות היחסים בין תהליכים פיזיקליים והתנסות ככל העולה על רוחנו. אף על פי כן, לכולנו יש גישה למקור שופע של נתונים במקרה של כל אחד מאיתנו. ניתן להסיק על החוקיות החשובה המתקיימת ביחס בין התנסות ועיבוד מתוך שיקולים הנוגעים להתנסות של כל אחד מאיתנו. קיימים גם מקורות עקיפים טובים של נתונים שעולים מתוך מקרים שניתן לחזות בהם, כמו למשל כשמישהו מסתמך על הדיווח הלשוני של סובייקט כאינדיקציה להתנסות. לשיטות אלה יש את המגבלות שלהן, אך יש לנו די והותר נתונים כדי להתחיל עם התיאוריה.

גם ניתוח פילוסופי הוא יעיל כדי להביאנו על גמולנו בעבור הנתונים שיש לנו. ניתוח שכזה יכול להניב מספר עקרונות הקשורים לתודעה והכרה, תוך שהוא מציב גבולות ברורים לצורתה הכוללת של התיאוריה. שיטת הניסויים המחשבתיים יכולה אף היא להניב שכר הולם, כפי שעוד נראה. ולבסוף, העובדה שאנו מחפשים תיאוריה פונדמנטלית פירושה הוא שאנו יכולים ליישם כללים לא-אמפיריים כגון פשטות והומוגניות בהבניית התיאוריה. עלינו לשאוף לארגן את המידע שיש לנו באופן סיסטמתי, להרחיבו עד כמה שאפשר בעזרת ניתוח זהיר, ואז ליצור היקש עבור התיאוריה הפשוטה ביותר שאפשר, אשר מסבירה את הנתונים ומהווה מועמדת סבירה להשתתפות בריהוט הפונדמנטלי של העולם.

לתיאוריות שכאלה תמיד יחדור איזשהו אלמנט של ספקולציה שאינו נוכח בתיאוריות מדעיות אחרות, וזאת כיוון שאי אפשר לערוך ניסויי-מבחן בין-סובייקטיביים שניתן להסתמך עליהם באופן מדעי. עם זאת, אנחנו בהחלט יכולים להבנות תיאוריות שמתיישבות עם הנתונים שיש לנו ולערוך השוואות ביניהן כדי לדעת איך להעריך אותן. אפילו בהיעדר תצפיות בין-סובייקטיביות, ישנם קריטריונים רבים שניתן להשתמש בהם כדי לדעת איך להעריך תיאוריות מסוג זה: פשטות, קוהרנטיות פנימית, קוהרנטיות ביחס לתיאוריות אחרות בתחומים אחרים, יכולת לייצר תכונות של ההתנסות המוכרות לנו מהמקרה הפרטי שלנו, ואף התאמה כללית לתכתיבי השכל הבריא. אולם גם כאשר כל ההגבלות האלה תיושמנה, עדיין ייוותרו דברים חשובים שלא הוגדרו, אבל ביכולתנו לפחות לפתח תיאוריות שתהיינה מועמדות סבירות. רק כאשר נפתח תיאוריות מתחרות, נהייה מסוגלים לדעת את ערכן.

תיאוריה לא-רדוקטיבית של התודעה מורכבת ממספר עקרונות פסיכופיזיים, כלומר עקרונות שמחברים את מאפייני התהליך הפיזיקלי עם מאפייני ההתנסות. נוכל לחשוב על עקרונות אלו כמגלמים את האופן שהתנסות נוצרת מתוך הדבר הפיזיקלי. עקרונות אלו הרי אמורים לומר לנו אילו מערכות פיזיקליות מקושרות להתנסות, ועבור המערכות המקושרות להתנסות, אילו תכונות פיזיקליות רלבנטיות להיווצרות ההתנסות, ואיזה סוג התנסות יוצרת כל מערכת פיזיקלית. אמנם מדובר בעבודה בסדר גודל רציני, אך אין סיבה שלא נתחיל בה.

בהמשך אציג את העקרונות הפסיכופיזיים האפשריים שעשויים להיכנס לתיאוריית תודעה. השניים הראשונים הם *עקרונות לא-בסיסיים* – כלומר קשרים סיסטמתיים בין עיבוד והתנסות ברמה יחסית גבוהה. עקרונות אלה יכולים לשחק תפקיד חשוב בפיתוח והגבלה של תיאוריית תודעה, אך הם אינם נבחרים בשלב שהוא פונדמנטלי דיו כדי לשמש כחוקים בסיסיים באמת. את העיקרון האחרון אציע כעיקרון *בסיס, אשר יהווה* את אבן הפינה של תיאוריית תודעה פונדמנטלית. עיקרון אחרון זה הנו ספקולטיבי במיוחד, אך הוא הספקולציה הנדרשת אם ברצוננו שתהיה לנו אי פעם תיאוריית תודעה. אוכל להציג כאן את העקרונות הללו רק בקצרה; אני מספק טיעונים עבורם ביריעה רחבה הרבה יותר בספרי "המיינד המודע"[[5]](#footnote-5).

**1**. עקרון הקוהרנטיות המבנית

זהו עיקרון של קוהרנטיות בין מבנה התודעה למבנה המודעות. במילה "מודעות" השתמשתי קודם כדי להתייחס לתופעות תפקודיות שונות המזוהות עם התודעה. כעת אני משתמש בה כדי להתייחס לתהליך ספציפי יותר בהצבת היסודות הקוגניטיביים של ההתנסות. את תכני המודעות יש להבין כעת כתכני המידע הנגישים למערכות מרכזיות ושנעשה בהם שימוש רחב כדי לשלוט בהתנהגות. אם לנסח זאת בקצרה, אנו יכולים לחשוב על המודעות כעל נגישות ישירה לשליטה גלובלית. תחילה ניתן לומר שתכני המודעות הם התכנים אשר נגישים באופן ישיר ושקיימת היכולת לדווח עליהם, לפחות במערכת המשתמשת בשפה[[6]](#footnote-6).

המודעות היא מונח תפקודי לחלוטין, אך למרות זאת היא קרובה מאוד להתנסות מודעת. ברוב המקרים, היכן שאנו מוצאים תודעה, אנו מוצאים מודעות. היכן שקיימת התנסות מודעת קיים איזשהו מידע מתאים במערכת הקוגניטיבית אשר נגיש עבור השליטה בהתנהגות ועבור הדיווח הלשוני. לעומת זאת, נדמה שברגע שמידע נגיש לדיווח ולשליטה כללית, ישנה התנסות מודעת בהתאם. לפיכך, יש התאמה ישירה בין תודעה ומודעות.

את ההתאמה ניתן להראות בדרכים נוספות. העובדה שלהתנסות יש מבנה מורכב היא עובדה מרכזית. לשדה הויזואלי יש למשל גאומטריה מורכבת. ישנם גם יחסי דמיון ושוני בין התנסויות, וישנם יחסים הנוגעים למשל לאינטנסיביות היחסית. ניתן לאפיין ולפרק כל התנסות של סובייקט לפחות באופן חלקי במונחים של התכונות המבניות הללו: יחסי דמיון ושוני, מיקום הנתפס בחושים, אינטנסיביות יחסית, מבנה גאומטרי, וכדומה. זוהי גם עובדה מרכזית שלכל אחד מן המאפיינים המבניים הללו יש מאפיין מתאים במבנה עיבוד המידע של המודעות.

הבה ניקח לדוגמה את ההתנסות בצבעים. לכל הבחנה בין התנסויות בצבע ישנה הבחנה מתאימה בעיבוד. הצבעים שאנו מתנסים בהם ושהנם שונים מבחינה פנומנלית יוצרים חלל תלת-ממדי מורכב, אשר משתנה לפי הגוון, הרוויה והבהיקות. ניתן לחלץ את התכונות של חלל זה מתוך עניינים הנוגעים לעיבוד מידע. בחינה של מערכות ויזואליות מראה שצורות גל של אור מובחנות ומנותחות לאורם של שלושה צירים, והמידע התלת-ממדי הזה הוא זה שרלבנטי לעיבוד מאוחר יותר. משום כך, המבנה התלת-ממדי של חלל צבע פנומנלי מתאים ישירות למידע שניתן לדווח עליו ולכן להבחנה שמוצגת במבנה העיבוד.

ניתן לומר במילים פשוטות יותר שהמבנה הגאומטרי של השדה הויזואלי מתגלה באופן ישיר במבנה שניתן לשחזר מתוך עיבוד ויזואלי. כל יחס גאומטרי מתאים למשהו שאפשר לדווח עליו והא לפיכך מיוצג באופן קוגניטיבי. לו היה ניתן לנו רק הסיפור אודות עיבוד המידע במערכת הויזואלית והקוגניטיבית של סובייקט, לא היינו יכולים לחזות באופן ישיר בהתנסויות הויזואליות האלו של זה שפועל, אולם למרות זאת היינו יכולים להקיש את התכונות המבניות האלו של ההתנסות.

באופן כללי, כל מידע שמתנסים בו באופן מודע ייוצג גם באופן קוגניטיבי. המבנה המגורען-דק של השדה החזותי יתאים לכמה מבנים מגורענים-דק בעיבוד חזותי. הדבר נכון גם להתנסויות במודוסים אחרים ואף להתנסויות לא-סנסוריות. לדימויים מנטליים פנימיים יש תכונות גאומטריות המיוצגות בעיבוד המידע. אפילו לרגשות יש תכונות מבניות, כמו אינטנסיביות יחסית, אשר מתאימות באופן ישיר לתכונה מבנית מסוימת של עיבוד: היכן שיש אינטנסיביות גבוהה יותר, אנו מוצאים השפעה גבוהה יותר על עיבוד מאוחר. באופן כללי, בדיוק מפני שהתכונות המבניות של התנסות נגישות וניתן לדווח עליהן, תכונות אלו ייוצגו באופן ישיר במבנה המודעות.

האיזומורפיזם הזה בין מבני התודעה והמודעות הוא בדיוק מה שמכונן את עקרון הקוהרנטיות המבנית. עיקרון זה מבוסס על העובדה החשובה שלמרות שתהליכים קוגניטיביים אינם מובילים מבחינה מושגית לעובדות הנוגעות להתנסות מודעת, תודעה ומודעות אינן מרחפות באופן חופשי אחת לצד השנייה אלא מחוברות באופן הדוק זו לזו.

לעיקרון זה יש את מגבלותיו. הוא מאפשר לנו לשחזר תכונות מבניות של התנסות מתוך תכונות עיבוד המידע, אך לא כל תכונות ההתנסות הן תכונות מבניות. ישנן תכונות מבניות, כגון טבעה האינהרנטי של תחושת האדום, שלא ניתן לתופסן לגמרי בתיאור מבני. עצם האפשרות להבין תרחישים של הספקטרום ההפוך, כלומר התנסויות של אדום וירוק שמתהפכות על אף שכל התכונות המבניות נותרות על כנן, מראה שהתכונות המבניות מגבילות את ההתנסות מבלי להיות ההתנסות עצמה. אף על פי כן, עצם העובדה שאנו מרגישים מחויבים לעזוב את התכונות המבניות ללא שינוי, ברגע שאנו מדמיינים התנסויות הפוכות בין מערכות זהות מבחינה תפקודית, מראה כמה מרכזי עקרון הקוהרנטיות המבנית עבור התפיסה שלנו את חיינו המנטליים. זה אינו עיקרון הכרחי מבחינה לוגית, שכן ככלות הכול ביכולתנו לדמיין את כל עיבוד המידע מתרחש ללא כל התנסות, אך יחד עם זאת, זו הגבלה חזקה וכללית החלה על הקשר הפסיכופיזיקלי.

עקרון הקוהרנטיות המבנית מאפשר הסבר יעיל מאוד ולא-ישיר של התנסות במונחים של תהליך פיזיקלי. לדוגמה, אנו יכולים להשתמש בעובדות הנוגעות לעיבוד עצבי של מידע ויזואלי כדי להסביר באופן עקיף את מבנה חלל הצבעים. העובדות הנוגעות לעיבוד עצבי יכולות לחשוף את מבנה המודעות ולהסבירו; אם ניקח את עקרון הקוהרנטיות כמובן מאליו, מבנה ההתנסות יוסבר גם הוא. חקירה אמפירית עשויה אף לגרום לנו להבין טוב יותר את מבנה המודעות של העטלף, ולשפוך בעקיפין אור על שאלתו המרתקת של נייגל: איך זה להיות עטלף? עיקרון זה מספק פרשנות טבעית למחקרים רבים בתחום של הסבר התודעה (למשל Clarck 1992, Hardin 1992 על צבעים, Akins 1993 על עטלפים), על אף שבדרך כלל טוענים שלא ניתן לומר דברים מפורשים בנושא. זה כל כך מוכר עד שכמעט כל אחד לוקח זאת כמובן מאליו וזה אלמנט בסיסי בהסבר הקוגניטיבי של התודעה.

הקוהרנטיות בין תודעה ומודעות יכולה גם לאפשר פרשנות טבעית למחקרים במדעי המוח השואפים לבודד את המתאמים העצביים של התודעה. על הפרשנות הזו אני מוסיף לפרט בפרק 4.

**2**. עקרון ההתמדה הארגונית

עיקרון זה קובע שלשתי מערכות כלשהן בעלות אותו הארגון התפקודי המגורען-דק יהיו התנסויות זהות מבחינת איכותן. אם השבלונות הסיבתיות של הארגון העצבי היו משוכפלות בסיליקון, למשל עם שבב סיליקון עבור כל נוירון ואותן שבלונות של אינטראקציה, אז היו נוצרות אותן ההתנסויות. על פי עיקרון זה, מה שחשוב להיווצרות התנסות אינו התבנית הפיזיקלית הספציפית של מערכת כלשהי אלא השבלונה המופשטת של אינטראקציה סיבתית בין רכיביה. עיקרון זה כמובן שנוי במחלוקת. חוקרים אחדים (למשל Searle 1980) חשבו שתודעה קשורה לביולוגיה מסוימת, כך שסיליקון הדומה במבנהו לאדם אינו אמור להיות מודע. אולם אני מאמין שניתן לספק תמיכה חשובה לעיקרון זה באמצעות ניתוח ניסויים מחשבתיים.

בקצרה מאוד: הבה נניח (לשם הטיעון עד אבסורד) שהעיקרון שגוי ושהיו יכולות להתקיים שתי מערכות בעלות מבנה דומה מבחינה תפקודית, אך בעלות התנסויות שונות. אולי רק אחת מהמערכות מודעת, או שאולי שתיהן מודעות, אך יש להן התנסויות שונות. לשם הסקיצה, הבה נאמר שמערכת אחת עשויה מנוירונים והאחרת מסיליקון, ושהאחת מתנסה באדום היכן שהשנייה בכחול. שתי המערכות הנן בעלות אותו הארגון כך שאנו יכולים לדמיין טרנספורמציה של האחת אל השנייה באופן הדרגתי, אולי על ידי החלפת כל נוירון בשבב סיליקון עם אותו התפקוד המקומי. בדרך זו אנו משיגים ספקטרום של מקרים סבירים, כל אחד עם אותו הארגון אך עם הרכב פיזיקלי מעט שונה והתנסויות מעט שונות.

לאורך הספקטרום הזה חייבות להיות שתי מערכות, א ו-ב, שמהן נעביר מזו לזו פחות מעשירית המערכת, אל זהו החלק שבו ההתנסויות שונות. שתי המערכות האלו זהות מבחינה פיזיקלית, מלבד העובדה שמעגל עצבי קטן ב-א הוחלף על ידי מעגל סיליקון ב-ב.

השלב המכריע בניסוי המחשבתי הוא ליטול את המעגל העצבי הרלבנטי ב-א ולשתול לצדו שבב סיליקון בעל צורה דומה, אף עם מפסק ביניהם. מה יקרה אם נפעיל את המפסק? ההיפותזה היא שההתנסויות המודעות של המערכת ישתנו, למשל מאדום לכחול. זה נובע מהעובדה שהמערכת לאחר השינוי היא במהותה גרסה של ב, בעוד שלפני השינוי היא מהווה רק את א.

עם זאת, בהינתן ההנחות הללו, למערכת אין שום דרך לשים לב לשינויים. הארגון הסיבתי שלה נותר קבוע, כך שכל מצביה התפקודיים ואופני התנהגותה נשמרו כפי שהם. ככל הנוגע למערכת, שום דבר מיוחד לא קרה. אין טעם למחשבה "הממ! קרה משהו מוזר!" בדרך כלל, המבנה של כל מחשבה שכזו חייב להתגלות בעיבוד, אך מבנה העיבוד נותר פה זהה. אם הייתה קיימת כזו מחשבה, היא וודאי הייתה מתקיימת מחוץ למערכת, חסרת כל יכולת להשפיע על עיבוד מאוחר יותר. (אם היא הייתה משפיעה על עיבוד מאוחר יותר, המערכות היו שונות מבחינה תפקודית, בניגוד להיפותזה). נוכל להפעיל את המפסק מספר פעמים, כך שהתנסויות באדום וכחול ירקדו הנה והנה אל מול "עיניה הפנימיות" של המערכת. לפי ההיפותזה, המערכת אינה יכולה להבחין ב"קוואליה המרקדת" הזו.

אני מחשיב זאת כהפרכה מוחלטת של ההנחה המוקדמת על דרך האבסורד.[[7]](#footnote-7) זו הרי עובדה ידועה שאנו מכירים בעצמנו: כל פעם שההתנסות משתנה באופן משמעותי ואנו שמים לב, אנחנו מסוגלים להבחין בשינוי; אם זה לא היה כך, היינו מובלים לאפשרות הספקנית שההתנסויות שלנו מרקדות לנו מול העיניים כל הזמן. להיפותזה זו יש אותו המעמד כמו לזו שאומרת שהעולם נברא לפני חמש דקות: אולי היא קוהרנטית מבחינה לוגית, אך היא אינה סבירה. בהינתן ההנחה הסבירה בהחלט שהשינויים בהתנסות מתאימים לשינויים בעיבוד המידע, אנחנו מובלים למסקנה שההיפותזה המקורית אינה אפשרית ושלכל שתי מערכות הדומות מבחינה תפקודית ישנו אותו סוג התנסויות. אם לנסח זאת באופן טכני, ההיפותזה הפילוסופית של "קווליה נעדרת" ו"קווליה הפוכה" היא אולי אפשרית מבחינה לוגית, אך מבחינה אמפירית ונומולוגית היא אינה אפשרית.

(יהיו את אלה שיעלו את החשש שסיליקון שדומה למערכת עצבית אינו אפשרי מסיבות טכניות. השאלה הזו פתוחה. עקרון ההתמדה אומר רק שאם מבנה דומה אפשרי, אז יהיה לו אותו סוג של התנסות מודעת.)

יש להוסיף ולומר כאן עוד כמה דברים, אך הבסיס החשוב כבר ניתן. הניסוי המחשבתי מתבסס על עובדות ידועות אודות הקוהרנטיות בין תודעה ועיבוד קוגניטיבי כדי להסיק מסקנה רצינית אודות היחס בין מבנה פיזיקלי והתנסות. אם הטיעון נכון, אנו יודעים שהתכונות הפיזיקליות היחידות שרלבנטיות להיווצרות ההתנסות הן תכונות ארגוניות. עניין זה פועל כמגבלה חזקה נוספת על תיאוריית התודעה.

**3**. תיאוריית המידע בעלת ההיבט הכפול

שני העקרונות הקודמים הנם עקרונות לא-בסיסיים. הם כוללים מונחים ברמה גבוהה כגון "מודעות" ו"ארגון" ולכן הם נמצאים ברמה הלא נכונה כדי לכונן את החוקים הפונדמנטליים בתיאוריית תודעה. יחד עם זאת, הם פועלים כמגבלות רציניות. מה שעוד זקוקים לו זה עקרונות בסיסיים שמתאימים למגבלות הללו ושעשויים להסבירן.

העיקרון הבסיסי שאני מציע כולל בעיקר את מושג המידע. אני תופס את רעיון המידע פחות או יותר כמו שאנון (1948). היכן שיש מידע, ישנם מצבי מידע המגולמים בחלל מידע. לחלל המידע יש מבנה בסיסי של יחסי שוני בין האלמנטים שבו, המאפיינים את הדרכים שבהן אלמנטים שונים בחלל דומים או שונים, גם ברמה מורכבת. חלל מידע הוא אובייקט מופשט, אך לפי שאנון ביכולתנו לראות מידע כמתגלם באופן פיזיקלי, כאשר יש חלל של מצבים פיזיקליים שונים, אשר את ההבדלים ביניהם ניתן לשדר בכמה ערוצים סיבתיים. את המצבים שניתן לשדר אפשר לראות כמכוננים בעצמם חלל מידע. אם ללוות ביטוי מבייטסון [Bateson] (1972), מידע פיזיקלי הוא שוני שמשנה [a difference, that makes a difference].

העיקרון בעל ההיבט הכפול נובע מכך שניתן לערוך תצפיות ולראות שיש דמיון צורני ישיר בין מידע מסוים המגולם באופן פיזיקלי לבין חללי מידע מסוימים פנומנליים (או התנסותיים). מאותן תצפיות שנכללו בעיקרון הקוהרנטיות המבנית אנו יכולים להסיק שלהבדלים בין המצבים הפנומנליים יש מבנה שמתאים ישירות להבדלים המגולמים בתהליכים פיזיקליים; ובמיוחד, להבדלים שיוצרים הבדלים דרך ערוצים סיבתיים מסוימים הנכללים בנגישות כללית ושליטה. הווה אומר, ביכולתנו למצוא חלל מידע מופשט כלשהו המגולם בעיבוד פיזיקלי ובהתנסות מודעת.

עניין זה מוביל להיפותזה טבעית: שלמידע (או לפחות למידע מסוים) יש שני היבטים, היבט פיזיקלי והיבט פנומנלי. לעניין זה יש מעמד של עיקרון יסוד שעשוי להוות את הבסיס להיווצרות ההתנסות ולהסבירה מתוך משהו פיזיקלי. ההתנסות נוצרת בעקבות הסטטוס שלה כהיבט אחד של מידע, כאשר ההיבט האחר מגולם בעיבוד פיזיקלי.

מספר מחשבות מסוגלות לתמוך בעיקרון זה, וביכולתי להביאן כאן רק בקצרה. ראשית, המחשבה על השינויים הפיזיקליים המתאימים לשינויים בהתנסות המודעת מגלה ששינויים כאלה תמיד רלבנטיים עקב תפקידם בכינון שינויי מידע – כלומר הבדלים בחלל מצבים מופשט, אשר מחולקים לפי ההבדלים הסיבתיים שלהם על פי מסלולים מסוימים. שנית, אם עקרון ההתמדה הארגונית נכון, אז אנו צריכים למצוא תכונה ארגונית פונדמנטלית כלשהי שצריך לחבר אותה להתנסות, ומידע הוא תכונה ארגונית מובהקת. שלישית, עיקרון זה מאפשר להסביר את עקרון הקוהרנטיות המבנית במונחים של המבנה הנוכח בחללי מידע. רביעית, הניתוח של ההסבר הקוגניטיבי עבור שיפוטינו וטענותינו בעניין ההתנסות המודעת – שיפוטים שניתן להסביר מבחינה תפקודית אך למרות זאת הם קשורים קשר עמוק להתנסות עצמה – מציע שההסבר כולל בעיקר את מצבי המידע המגולמים בעיבוד קוגניטיבי. מכאן נובע שתיאוריה המבוססת על מידע מאפשרת קוהרנטיות עמוקה בין ההסבר של ההתנסות לבין ההסבר של שיפוטינו וטענותינו אודותיה.

ווילר [Wheeler] (1990) הציע שהמידע הוא פונדמנטלי עבור הפיזיקה של היקום. לפי תורה זו, תורת ה"it from bit", ניתן לנסח את חוקי הפיזיקה במונחים של מידע ולהצביע על מצבים שונים המחוללים השפעות שונות מבלי לומר ממש מהם מצבים אלו. רק מיקומם בחלל המידע הוא מה שחשוב. לפיכך, מידע הוא מועמד טבעי לשחק תפקיד בתיאוריה פונדמנטלית של התודעה. אנו מובלים לתפיסת העולם שבה מידע הוא פונדמנטלי באמת ושבה יש לו שני היבטים חשובים, אחד שמתאים למאפיין הפיזיקלי של העולם, ואחד שמתאים למאפיין הפנומנלי.

כמובן שהעיקרון בעל ההיבט הכפול הוא ספקולטיבי לחלוטין ועמום במקצת, והוא מותיר מספר שאלות מפתח ללא מענה. שאלה מתבקשת היא האם לכל מידע יש היבט פנומנלי. אפשרות אחת היא שאנו צריכים מגבלה נוספת לתיאוריה הפונדמנטלית שתראה פשוט לאיזה סוג של מידע יש היבט פנומנלי. האפשרות האחרת היא שאין כזו מגבלה. אם אין, אזי ההתנסות היא נפוצה הרבה יותר ממה שחשבנו, שכן ההתנסות היא בכל מקום. בתחילה, העניין זה הוא קונטרה-אינטואיטיבי, אך כשבוחנים זאת העמדה הופכת לסבירה ואלגנטית. היכן שיש פשוט עיבוד מידע, ישנה התנסות מורכבת. לעכבר יש מבנה פשוט יותר של עיבוד מידע מאשר לאדם ולכן יש לו בהתאם פחות התנסות; האם לתרמוסטט, שהוא מבנה עיבוד מידע פשוט ביותר, יש התנסות פשוטה ביותר? אם ההתנסות היא תכונה פונדמנטלית באמת, יהא זה אכן מפתיע אם היא תיווצר רק מדי פעם; רוב התכונות הפונדמנטליות מפוזרות בצורה יותר שווה. בכל מקרה, זו שאלה די פתוחה, אך אני חושב שהעמדה אינה כה בלתי סבירה כפי שהיא נדמית על פי רוב.

ברגע שמונח על השולחן קישור פונדמנטלי בין מידע והתנסות, הדלת פתוחה לספקולציה מטאפיזית יותר גרנדיוזית אודות טבע העולם. למשל, לרוב סבורים שהפיזיקה מאפיינת את יחידותיה הבסיסיות רק באופן חיצוני, על פי יחסיהן ליחידות אחרות, שהן בעצמן מאופיינות באופן חיצוני, וכן הלאה. לטבען הפנימי של יחידות פיזיקליות לא מתייחסים. יש הטוענים שתכונות פנימיות כאלה אינן קיימות, אך אז נותרים עם עולם שהוא זרם סיבות טהור (זרם טהור של מידע) ללא תכונות שהסיבתיות יכולה לחבר ביניהן. אם מודים בקיומן של תכונות פנימיות , הספקולציה הטבעית, בהתחשב במה שנאמר קודם, תהיה לטעון שהתכונות הפנימיות של מה שפיזיקלי – התכונות שהסיבתיות מחברת בסופו של דבר – הן בעצמן תכונות פנומנליות.[[8]](#footnote-8) אנחנו יכולים לומר שתכונות פנומנליות הן ההיבט הפנימי של מידע. עניין זה יכול לענות על חשש הנוגע לרלבנטיות הסיבתית של התנסות – חשש טבעי, בהינתן תמונה שבה התחום הפיזיקלי סגור מבחינה סיבתית ושבה התנסות היא תוספת כלשהי לתחום הפיזיקלי. ההשקפה המבוססת על מידע מאפשרת לנו להבין כיצד להתנסות עשוי להיות סוג מתוחכם של רלבנטיות סיבתית על בסיס מעמדה כטבע פנימי של דברים פיזיקליים. על מנת להבנות תיאוריה מדעית, כנראה שעדיף להתעלם מספקולציה מטאפיזית שכזו, אבל כשעוסקים בעניינים פילוסופיים זה די מתבקש.

**8. מסקנה**

התיאוריה שהצגתי היא ספקולטיבית, אך היא תיאוריה שעומדת למבחן. אני חושד שעקרון הקוהרנטיות המבנית ועקרון ההתמדה הארגונית יהיו רכיבים בכל תיאוריית תודעה משביעת רצון; המעמד של תיאוריית המידע בעלת ההיבט הכפול הוא פחות ברור. אכן, כרגע הוא יותר רעיון מאשר תיאוריה. כדי להשיג איזושהי הצלחה בנוגע להסברים, היא זקוקה לפירוט רחב יותר ולעיבוי לכדי צורה חזקה יותר. ובכל זאת, רפלקסיה בנוגע למה שסביר ולא סביר בה, ובנוגע להיכן היא עובדת והיכן היא כושלת, יכולה רק להוביל לתיאוריה טובה יותר.

רוב תיאוריות התודעה הקיימות מכחישות את התופעה, מסבירות משהו אחר, או מרוממות את הבעיה לכדי תעלומה נצחית. אני מקווה שהצלחתי להראות שניתן להתקדם עם הבעיה ברגע שלוקחים אותה ברצינות. כדי להתקדם עוד, עלינו לחקור עוד, ואנו זקוקים לתיאוריות יותר מלוטשות ולניתוחים יותר זהירים. הבעיה הקשה היא אכן בעיה קשה, אך אין סיבה להאמין שהיא תיוותר לעד בלתי פתירה.[[9]](#footnote-9)

1. רשימה זו אינה מושלמת מפני שרוב התופעות שברשימה קשורות באופן מסוים להתנסות, ועניין זה מעורר את הבעיה הקשה. קיימים סוגי התנסויות שונים הקשורים, למשל, לתשומת לב ושליטה מרצון. יש להבין את הרשימה ככזו שמפנה את תשומת הלב בעיקר להיבטים הפונקציונליים של התופעות האלו, ופחות להיבטים ההתנסותיים. יש לציין שהבעיות הקלות *אינן* מאופיינות כבעיות של האינטנציונליות. [↑](#footnote-ref-1)
2. לעיתים סבורים שהטיעונים שבעד חוסר ההסבר משתמשים במודל "דדוקטיבי-נומולוגי" מיושן של הסבר. לכן חשוב לציין שהפרק הזה מתייחס למודל של הסבר המתבסס על תפקודים ומנגנונים. מודל זה קשור קשר הדוק למודלים שנעשו פופולריים בפילוסופיה של המדע בשנים שעברו מאז שהתפרסם המאמר שעליו מבוסס הפרק הזה. [↑](#footnote-ref-2)
3. בתשובה למאמר שעליו מבוסס הפרק הזה, האמרוף ופנרוז (1996) מאשרים במפורש את הרעיון שהתנסות היא יסוד בסיסי בטבע. [↑](#footnote-ref-3)
4. גארט (2006) טוען שויטליסט אחד בשם נחמיה גרו סיפק טיעונים המתבססים על יכולתו של דבר-מה להיתפס באופן מושגי ואשר קשורים קשר הדוק לאלה הנוגעים לתודעה. אני חושב שאם בוחנים את זה מגלים שגם השקפתו של גרו מתיישבת עם הרעיון שמעלים את הרוח החיה באוב כדי להסביר את התפקודים. אלא שברגע שהרוח החיה עולה באוב בתפקיד זה, ניתן לתאר מערכות היפותטיות אחרות שמהן היא נעדרת. [↑](#footnote-ref-4)
5. את שלושת האלמנטים שיבואו בהמשך פיתחתי בפרקים 8-6 ב"מיינד המודע". הקוראים יכולים להרגיש חופשי לדלג על החומר הזה או רק לרפרף בו, הואיל והוא אינו חיוני לנרטיב של שאר הספר. [↑](#footnote-ref-5)
6. מושג המודעות הזה קשור קשר הדוק למושג "תודעת הגישה" של נד בלוק [Ned Block] (1995). למעשה, עיקרון זה מציע קוהרנטיות מסוימת בין המבנה של תודעה פנומנלית והמבנה של תודעה הגישה. את הנקודה הזו פיתחתי בפרשנות לבלוק (צ'אלמרס 1997א). את הקשר בין נגישות והתנסות אני בוחן גם בפרק 4. [↑](#footnote-ref-6)
7. אני עדיין חושב שההיפותזה הזו משונה, אך כעת אני נוטה לחשוב שהיא משהו אחר מהפרכה מוחלטת על דרך האבסורד. מחקרים בנושא העיוורון-לשינויים הרגילו אותנו לרעיון שלעתים איננו מבחינים בשינויים גדולים בתודעה. יש להודות ששינויים אלה מצויים מחוץ לתשומת הלב, והשינויים בתכני תשומת הלב שלא הבחנו בהם הם עוד יותר מוזרים, אך כנראה שזה לא ממש מוזר עד כדי כך שזה יישלל בכל הנסיבות. גם מוניזם ראסליאני (ראו פרק 5) מספק מודל טבעי שבו שינויים כאלה יכולים להתרחש. בספרי "המיינד המודע" הצעתי שטיעון ה"קוואליה המרקדת" הזה היה איכשהו חזק יותר מטיעון ה"קוואליה המתפוגגת" שניתן שם; כעת הייתי הופך את השיפוט הזה. [↑](#footnote-ref-7)
8. כדי למצוא עוד חומר בנושא ראו את הדיון על מוניזם מסוג F בפרק 5 ומוניזם ראסליאני בפרק 6. [↑](#footnote-ref-8)
9. קריאה נוספת: בבעיות התודעה דנו באופן מקיף בספרות הפילוסופית האחרונה. כדי לקבל הבהרות מושגיות נוספות אודות הבעיות השונות של התודעה, ראו: Block (1995), Nelkin (1993), Tye (1995). אלו שהדגישו את הסבר ההתנסות במונחים פיזיקליים כוללים, בין היתר, את Hodgson (1991), Jackson (1982), Levine (1983), Lockwood (1989), McGinn (1989), Nagel (1974), Seager (1991), Searle (1991), Strawson (1994), Velmans (1991). אלו שנקטו גישה רדוקטיבית כוללים את Churchland (1995), Clark (1992), Dennett (1991), Dretske (1995), Kirk (1994), Rosenthal (1997), Tye (1995). לא נעשו מאמצים רבים להבנות תיאוריות לא-רדוקטיביות מפורטות בספרות, אך כדי למצוא כמה מחשבות בכיוון ראו Hodgson (1991), Lockwood (1989). שני אוספים מצוינים של מאמרים בנושא התודעה הם Block, Flanagan, and Güzeldere (1997), Metzinger (1995). [↑](#footnote-ref-9)