



חוברת קריאה

פסיכולוגיה

עשרים ממצאים מפורסמים, ומה קרה כשניסו לחזור עליהם.

לא סיור בין תיאוריות אלא עשרים דברים שהפסיכולוגיה טענה על בני אדם — ומה נשאר מהם אחרי שהתחום הריץ על עצמו בדיקה. בכל יחידה שלושה שלבים: מה נטען, מה הראה המחקר המקורי, ומה קרה כשאחרים ניסו שוב. חלק מהממצאים קרסו, חלק התכווצו, וחלק דווקא יצאו מהבדיקה חזקים יותר — וההבחנה ביניהם היא הדבר החשוב ביותר שיש כאן ללמוד.

החוברת הזאת היא אותו חומר שנמצא באתר, ערוך לקריאה רצופה. כל יחידה עומדת בפני עצמה ואפשר לקרוא אותן בכל סדר, אבל הן מסודרות בסדר שיש בו היגיון — מי שקורא מן ההתחלה יראה קו מתפתח ולא רשימה.

בסוף כל יחידה יש סילבוס: מה לקרוא, לצפות ולהאזין כדי להעמיק. כל קישור שם נפתח ונבדק לפני שנכנס, וכתובתו המלאה מודפסת מתחתיו כדי שאפשר יהיה להקליד אותה גם מדף מודפס. היכן שיש ספרים, הם מופיעים בנפרד — שם, מחבר ושנה, בלי קישורי קנייה שמתיישנים.

היחידות שעוסקות במחלוקת מציגות את שני הצדדים בטיעון החזק ביותר שלהם, ואינן מכריעות ביניהם. זו בחירה מכוונת: המטרה היא שתוכלי להחזיק את הוויכוח בעצמך, לא לקבל ממני מסקנה מוכנה.

תוכן העניינים

תפיסה וקשב

01	הגורילה שאיש לא רואה	5
02	המוח כמנחש	8

זיכרון

03	הזיכרון אינו הקלטה	11
04	זיכרונות שווא ועדי ראייה	14
05	איך באמת לומדים	17

חשיבה והחלטות

06	שתי המערכות	20
07	מסגור ושנאת הפסד	23
08	ביטחון יתר וניבוי	26

שליטה עצמית

09	מבחן המרשמלו	29
10	דלדול האגו	32

התפתחות והתקשרות

11	פיאז'ה ומה שנשאר ממנו	35
12	התקשרות	38

פסיכולוגיה חברתית

13	מילגרם	41
14	כלא סטנפורד	44

47		אפקט המתבונן	15
50		תנוחות כוח	16

אישיות ואינטליגנציה

53		חמשת הגורמים ומיירס-בריגס	17
56		מנת משכל	18

המקצוע עצמו

59		מה באמת עובד בטיפול	19
62		משבר השחזור ומה הוא שינה	20

הגורילה שאיש לא רואה

1999 · מה קורה כשהקשב עסוק במשהו אחר

הניסוי פשוט עד כדי עלבון. צופים מתבקשים לספור כמה פעמים מוסרים את הכדור שחקנים בחולצות לבנות בסרטון קצר. באמצע הסרטון נכנסת לפריים אישה בתחפושת גורילה, עוצרת במרכז, מכה על חזה ויוצאת. היא נמצאת שם תשע שניות. כמחצית מהצופים לא רואים אותה.

המספר המדויק שווה תשומת לב, כי הוא מצוטט לא נכון בכל מקום. במחקר של סיימנס וצ'ברים מ-1999 נותחו 192 צופים בשמונה תנאים שונים, ובסך הכול 46 אחוז פספסו את האירוע. אבל בתוך התנאים הטווח עצום: מאפס אחוז ועד תשעים ושניים. כמה קשה משימת הספירה, אם הסרטון שקוף או אטום, אם הדמות היא גורילה או אישה עם מטרייה — כל אלה מזיזים את התוצאה. "חמישים אחוז" אינו חוק טבע אלא ממוצע.

מה שהופך את הממצא לחזק הוא לא הסרטון אלא מה שקרה אחריו. ב-1998 עצרו סיימנס ולוין עוברי אורח ברחוב ושאלו אותם לדרך; באמצע השיחה חלפו ביניהם שני אנשים שנשאו דלת, ובמקום בן השיח הקודם הופיע אדם אחר — בגדים אחרים, קול אחר, הפרש גובה של כחמישה סנטימטרים. שבעה מתוך חמישה עשר שמו לב. ב-2013 הודבקה תמונת גורילה לתוך סריקת חזה, גדולה פי ארבעים ושמונה מהגוש שאותו חיפשו; מתוך עשרים וארבעה רדיולוגים מקצועיים, עשרים לא ראו אותה — ומעקב עיניים הראה ששנים עשר מהם הביטו בה ישירות, בממוצע חצי שנייה. ב-2021 חזרו על כך עם חמישים רדיולוגים וממצא קליני אמיתי: שני שלישים פספסו גוש שד גדול בזמן שחיפשו גושים בריאה, וכשביקשו מהם לחפש אותו במפורש — תשעים ושבעה אחוז מצאו. הפיקסלים לא השתנו. השאלה השתנתה.

שלוש טעויות נפוצות כדאי לא לעשות. הראשונה, שזה עניין של עין: מעקב העיניים מראה שאנשים מתמקדים ישירות בגורילה ולא רואים אותה, כלומר הצוואר בקבוק הוא הקשב ולא האופטיקה. השנייה, שאנשים חכמים או קשובים במיוחד כן יבחינו: מטא-אנליזה מ-2024 על שבעים וארבעה מדגמים בדקה זיכרון עבודה, אינטליגנציה ותכונות אישיות, וכל המתאמים קרובים לאפס. השלישית, שידיעה מחסנת: בגרסה שבה נבדקו אנשים שכבר הכירו את הסרטון, כולם ראו את הגורילה — ולא היטיבו יותר מהתמימים לזהות את שני השינויים האחרים שנעשו בסרטון באותו זמן.

מה כן התערער? לא הקיום, אלא הפרשנות. ב-2023 הראה מחקר על 4,493 נבדקים שעצמים **מהירים** נתפסים כמעט תמיד — כשהגורילה חוצה את הפריים במהירות, למעלה מתשעים אחוז רואים אותה. וב-2024 נבדקו כ-25 אלף איש, ומי שדיווחו שלא ראו כלום ניחשו בכל זאת נכון מעל לרמת המקרה את מיקום העצם ואת צבעו. כלומר ייתכן שהקשב אינו מוחק את התפיסה אלא מחליש אותה, ושחלק ממה שנספר כ"לא ראיתי" הוא בעצם "לא הייתי בטוח מספיק כדי לומר". אף אחד מהחוקרים האלה לא טוען שהתופעה מדומה. זה בדיוק ההבדל בין ממצא שקורס לממצא שמתחדד — והיכולת להבחין ביניהם היא מה שהפקולטה הזאת מנסה ללמד.

זה אחד המעטים בפקולטה הזאת שיצא מכל בדיקה חזק יותר — ולכן הוא המקום הנכון להתחיל בו.

שנה	1999
תחום	תפיסה וקשב
נושא	עיוורון קשבי

להעמיק

חיפוש ביוטיוב: ניסוי הגורילה הבלתי נראית

לצפייה

כדאי לצפות בסרטון הספירה לפני שקוראים את היחידה, אחרת הידיעה הורסת את הניסוי.

https://www.youtube.com/results?search_query=ניסוי הגורילה הבלתי נראית

הגוש שאיש לא חיפש, 2021

לקריאה

ההמשך הקליני: חמישים רדילוגים, שני שלישים פספסו גוש שד גדול באותה סריקה, וכשהתבקשו לחפש אותו — תשעים ושבעה אחוז מצאו. באנגלית.

<https://link.springer.com/article/10.3758/s13423-020-01826-4>

הגיע הזמן לשדרג את הגורילה

לקריאה

מסה בעברית שטוענת שהניסוי אינו מוכיח שאנחנו עיוורים אלא שהתפיסה מכוונת מטרה — בדיוק הוויכוח הפרשני שהיחידה מציגה.

<https://alaxon.co.il/article/הגיע-הזמן-לשדרג-את-הגורילה>

תקציב האנרגיה של המוח קבוע

לקריאה

כתבה בעברית על מחקרה של נילי לביה: הסבר אנרגטי לעיוורון הקשבי, והקשר לתאונות דרכים ולטעויות רפואיות.

<https://www.hayadan.org.il/the-energy-budget-of-the-brain-is-fixed-0508205>

המאמר המקורי: "גורילות בקרבנו", 1999

מקורות וארכיון

הטקסט המלא של סיימונס וצ'בריס, עם טבלת שיעורי הפספוס שנעים בין אפס לתשעים ושניים אחוז, וההסתייגות של המחברים עצמם. באנגלית.

https://labs.wsu.edu/attention-perception-performance/documents/2016/05/simons_chabris_99.pdf/

הגורילה בסריקת הסייטי, 2013

מקורות וארכיון

עשרים וארבעה רדיולוגים, גורילה גדולה פי ארבעים ושמונה מגוש ריאה, ומעקב עיניים שמראה שרובם הביטו בה. באנגלית.

<https://search.bwh.harvard.edu/new/pubs/DrewVoWolfe13.pdf>

ספרים

הגורילה הבלתי נראית

כריסטופר צ'בריס ודניאל סיימונס · 2010

שני החוקרים שערכו את הניסוי מרחיבים אותו לשש אשליות של אינטואיציה, בעברית בתרגום אורי שגיא, כנרת זמורה-ביתן 2012.

Inattention Blindness

1998 · Arien Mack & Irvin Rock

הספר שקדם לסרטון והראה את אותה תופעה בלי וידאו ובלי גורילה. לא תורגם לעברית.

Eye and Brain

1966 · Richard L. Gregory

הקלאסיקה על תפיסה כהסקה ולא כהקלטה, מועילה גם ליחידה הבאה. לא תורגם לעברית.

המוח כמנחש

2015 · השמלה, מקגר, והיכן נגמרת ההסקה

בפברואר 2015 עלתה לרשת תמונה של שמלה, וחצי מהאינטרנט ראה לבן וזהב וחצי ראה כחול ושחור. השמלה עצמה כחולה ושחורה; הפיקסלים בתמונה כחולים וחומים. במחקר שנערך מיד, על כ־1,400 איש, 57 אחוז דיווחו כחול ושחור, 30 אחוז לבן וזהב, ואחד עשר אחוז כחול וחום — כלומר רק מיעוט קטן דיווח על מה שנמצא בתמונה בפועל.

ההסבר המוביל הוא שכל צופה מניח, בלי לדעת, באיזה אור צולמה התמונה. מי שמניח אור יום כחלחל מנכה את הגוונים הקרים ורואה לבן וזהב; מי שמניח תאורת פנים צהובה מנכה את החמים ורואה כחול ושחור. כשמוסיפים לתמונה רמזי תאורה מפורשים, התפיסה מתהפכת. במחקר המשך על 13,417 נבדקים נמצא גם מתאם מוזר ועקבי: מי שמגדירים את עצמם כאנשי בוקר נוטים ללבן וזהב, ואנשי לילה לכחול ושחור — בדיוק כפי שהיינו מצפים אם ההנחה על התאורה נבנית מהאור שאדם רגיל לחיות בו. חשוב לומר שאין עדיין הסכמה מלאה על המנגנון; שלוש קבוצות פרסמו שלושה הסברים מתחרים באותו גיליון.

הדוגמה השנייה היא אפקט מקגר: כששומעים הברה אחת ורואים שפתיים מבטאות אחרת, רבים שומעים הברה שלישית שאיש לא הפיק. גם כאן שווה לדייק. במחקר על 360 נבדקים השתנתה הרגישות לאפקט בין אפס למאה אחוז בין אנשים, ושבעים ושבעה אחוז הגיבו הכול או כלום; מי שנבדק שוב שנה אחר כך קיבל כמעט בדיוק את אותה תוצאה. בסקירה שעברה על ארבעים שנות מחקר נמצא שהשיעורים המדווחים נעים בין עשרים לתשעים ושמונה אחוז, ושכ־15 אחוז מהמבוגרים אינם חווים את האפקט כלל. כלומר זו לא תופעה שקורית "לכולם", וגם לא מדד טוב לאינטגרציה בין חושים — היא אינה מנבאת כמה טוב אדם מבין דיבור ברעש.

והנה החלק שמראה למה הפקולטה הזאת קיימת. במשך עשורים לימדו שאשליית מולר-לייר — שני קווים באותו אורך שנראים שונים בגלל החצים בקצותיהם — פועלת חזק על מערביים ולא על מי שגדלו מחוץ ל"עולם נגרי" מלא זוויות ישרות. המחקר שעליו הכול נשען, משנות השישים, נבחן מחדש ב־2025 ונמצא מלא סתירות פנימיות: קבוצות שלפי ההשערה היו אמורות להיות רגישות מאוד נמצאו מהעמידות ביותר, המדידה נעשתה דרך מתורגמנים שהחוקרים עצמם כתבו שאינם סומכים עליהם, ואחד המבצעים תיאר שפיתח ציפיות חזקות ונטייה לתקן תשובות. בינתיים האשליה נמצאה בדגים, בציפורים, בסוסים, בנמלים — ובילדים שנולדו עיוורים וקיבלו ראייה בניתוח: מאה אחוז מהם נפלו בה, לפני שראו בחייהם דלת או פינת חדר. האשליה שרדה. ההסבר התרבותי שלה קרס.

מכאן גם האזהרה. המשפט "המוח מייצר הזיה מבוקרת" הוא סיסמה יפה, ומסגרת ה"מוח המנבא" שמאחוריה היא מסגרת מחקר מעניינת — לא ממצא. ובכיוון ההפוך: שורה ארוכה של מחקרים שטענו שרצון או צורך

משנים ממש את מה שרואים — שילדים עניים רואים מטבעות גדולים יותר, שגבעה נראית תלולה יותר עם תרמיל כבד — פורקו ב־2016 והתבררו כמדידות של שיפוט ושל ציפיות הנבדק, לא של תפיסה. התפיסה אכן בונה ולא מקליטה. היא פשוט בונה מדויק הרבה יותר ממה שהסיסמה מרמזת.

למה זה ברשימה

כי השמלה ההיא הוכיחה לכמה מיליוני אנשים בבת אחת שמה שהם רואים אינו מה שיש שם, והשאלה מה בדיוק זה מוכיח עדיין פתוחה.

שנה	2015
תחום	תפיסה וקשב
נושא	תפיסה כהסקה

להעמיק

אניל סת': המוח שלך הוזה את המציאות

לצפייה

הרצאה בת שבע עשרה דקות עם תרגום לעברית — הניסוח הפופולרי החזק ביותר של הרעיון, ובדיוק זה שהיחידה מבקשת גם לצטט וגם לסייג.

https://www.ted.com/talks/anil_seth_your_brain_hallucinates_your_conscious_reality?language=he

האם תרבות משנה את מה שאנחנו רואים

לקריאה

הסבר בהיר על השערת "העולם הנגרי" מ־1966 ועל המחקר שפירק אותה, כולל הראיות מחיות ומילדים שנולדו עיוורים. באנגלית.

<https://www.scientificamerican.com/article/does-culture-change-visual-perception-debunking-the-carpentered-world/>

אין נביא במוחנו

לקריאה

מסה בעברית על ההיסטוריה של מטפורת המוח המנבא — ולמה כל דור מתאר את המוח לפי הטכנולוגיה של זמנו.

<https://alaxon.co.il/article/אין-נביא-במוחנו>

האם התפיסה החזותית "מערבית"?

מקורות וארכיון

המאמר המלא: הביקורת המתודולוגית על המחקר המקורי, והראיות לכך שהאשליה עתיקה ומולדת. באנגלית.

https://perception.jhu.edu/files/PDFs/25_MullerLyer/AmirFirestone_MullerLyer_2025_PsychReview.pdf

השמלה: הנחות סמויות על תאורה

מקורות וארכיון

המחקר על 13,417 נבדקים, כולל ממצא העירוני מול הינשוף וההסתייגויות של המחבר עצמו. באנגלית.

<https://jov.arvojournals.org/article.aspx?articleid=2617976>

ארבעים שנה לאפקט מקגרק

מקורות וארכיון

סקירה ביקורתית שמראה עד כמה האפקט תלוי אדם ותלוי גירוי, והאזהרה מלהשתמש בו כמדד לאינטגרציה בין חושים. באנגלית.

<http://brain.phgy.queensu.ca/pare/assets/Alsius2017.pdf>

ספרים

Being You

2021 · Anil Seth

הגרסה המורחבת של ההרצאה: תודעה, תפיסה כניחוש מבוקר, והיכן המסגרת הזאת עדיין השערה. לא אותר תרגום עברי.

The Experience Machine

2023 · Andy Clark

הניסוח הפופולרי של תיאוריית המוח המנבא מאת מי שפיתח אותה. לא אותר תרגום עברי.

Eye and Brain

1966 · Richard L. Gregory

הספר שניסח ראשון את התפיסה כהשערה שהמוח בודק, עשרות שנים לפני שהמסגרת קיבלה שם חדש. לא אותר תרגום עברי.

הזיכרון אינו הקלטה

1974 · מה עושה מילה אחת לזיכרון של תאונה

בניסוי מ-1974 הראו לארבעים וחמישה סטודנטים סרט של תאונת דרכים ושאלו אותם באיזו מהירות נסעו המכוניות. לכל קבוצה נשאלה אותה שאלה במילה אחרת: התנגשו, פגעו, נגעו, נחבטו, התרסקו. מי שנשאל "התרסקו" אמד ארבעים ואחת מייל לשעה; מי שנשאל "נגעו" אמד שלושים ושתיים. אותו סרט בדיוק. בניסוי שני, שבוע אחר כך, נשאלו מאה וחמישים נבדקים אם ראו זכוכית שבורה — ולא הייתה זכוכית שבורה בסרט. מבין מי שנשאלו קודם במילה "התרסקו", שישה עשר מתוך חמישים "ראו" אותה, לעומת שבעה מקבוצת "פגעו".

מאחורי הניסוי הזה עומד רעיון ישן יותר. כבר ב-1932 הראה ברטלט שאנשים ששמעו סיפור מתרבות זרה סיפרו אותו מחדש כשהוא מעוצב לפי הסכמות שלהם. שיטת העבודה שלו הייתה רופפת מאוד לפי כל קנה מידה מודרני — בלי סטטיסטיקה, בלי גודל מדגם ברור — ולכן שווה לדעת מה קרה כשבדקו אותו כראוי ב-1999: הסיפור פורק ל-42 יחידות תוכן ונמדד על פני חצי שנה. הדיוק צנח כמעט לאפס, ומספר העיוותים הגדולים **לא** ירד. הם לא נשכחו. הם התקבעו. המסקנה של ברטלט שרדה בדיוק מפני שמישהו טרח לבדוק אותה.

אפקט המידע המטעה — העובדה שמה שנאמר לך אחרי אירוע נכנס לזיכרון של האירוע — הוא אחד הממצאים החסונים בפסיכולוגיה. מטא-אנליזה מ-2014 סוכמת 155 גדלי אפקט, והוא משמש כבסיס מדידה יציב בעשרות ניסויים. גם דרך ההפחתה שלו ידועה: אזהרה שניתנת **אחרי** החשיפה למידע המטעה חותכת את האפקט לכשליש עד מחצית — לא מאפסת אותו. ובכל זאת אסור להגזים גם כאן: בניסיון שחוזר מ-2017 בהקשר של ספורט, אפקט הפועל עצמו לא הופיע בגרסה הכתובה, והופיע בגרסה משודרת בכיוון הפוך. הדירוג בין "התרסקו" ל"נגעו" עשוי להיות תלוי תחום.

החלק שמכה חזק ביותר הוא דווקא זה שנוגע לזיכרונות שנראים חקוקים. הרעיון של "זיכרון פלאש" — שרגעים דרמטיים נצרכים בדיוק צילומי — נבדק בקונסורציום שעקב אחרי 391 איש בשבע ערים אחרי ה-11 בספטמבר, במדידות אחרי שבוע, אחרי שנה ואחרי שלוש. העקביות ירדה ל-63 אחוז בשנה הראשונה ול-57 אחוז אחרי שלוש, בזמן שהביטחון נשאר גבוה, בסביבות 4.3 מתוך 5. במעקב אחרי עשר שנים נמצא הדבר שקשה להתאושש ממנו: המתאם בין העקביות לבין הביטחון היה 0.07 — כלומר אפס. אדם שבטוח לחלוטין ואדם שמהסס נכונים באותה מידה. ועוד: 43 אחוז מאי ההתאמות במדידה הרביעית היו חזרה על אותן שגיאות מהמדידה השלישית. הזיכרון לא סתם נשחק; הוא נעול על הגרסה החדשה.

מה שאסור להסיק מכל זה הוא שהזיכרון בדוי. אחרי עשר שנים העקביות היא כ-59 אחוז, לא אפס, והדיוק לגבי עובדות האירוע עצמו התייצב סביב 77 אחוז. הזיכרון עובד — הוא פשוט לא עובד כמו מצלמה, והביטחון

הפנימי שלנו אינו מד שאפשר לסמוך עליו. ליחידה הבאה יש לזה השלכה מעשית מאוד, כי המערכת שבה מסתמכים על ביטחון של עדים היא מערכת המשפט.

למה זה ברשימה

כי הביטחון שבו את זוכרת משהו אינו מעיד על הדיוק שלו — וזה נמדד, במספרים, על פני עשר שנים.

שנה	1974
תחום	זיכרון
נושא	שחזור זיכרון

להעמיק

הטיות זיכרון

לקריאה

פתיחה עברית נגישה על זיכרונות שווא ועל הדרך שבה מידע שנכנס אחרי האירוע משנה אותו.

<https://davidson.org.il/read-experience/firefly/memory-biases/>

לופטוס ופאלמר, ניסוי תאונת הדרכים

לקריאה

כל המספרים של שני הניסויים המקוריים: אומדני המהירות לכל פועל, ואחוזי מי ש"ראה" זכוכית שבורה. באנגלית.

<https://www.simplypsychology.org/loftus-palmer.html>

זיכרון ארוך טווח לאירועי 11 בספטמבר

מקורות וארכיון

מחקר האורך על שבע ערים: 63 אחוז עקביות אחרי שנה מול ביטחון של 4.41 מתוך 5. באנגלית.

<https://www.memlab.psychol.cam.ac.uk/pubs/Hirst2009 JPEG.pdf>

מעקב עשר שנים

מקורות וארכיון

ההמשך שבו המתאם בין ביטחון לעקביות מתאפס, והשגיאות מתקבעות במקום להיתקן. באנגלית.

<https://www.memlab.psychol.cam.ac.uk/pubs/Hirst2015 JPEG.pdf>

האם אפשר לשחזר את ניסויי ברטלט

מקורות וארכיון

השכפול המבוקר שאישר את מסקנתו והראה בדיוק כמה השיטה המקורית שלו הייתה רופפת. באנגלית.

http://psychnet.wustl.edu/memory/wp-content/uploads/2018/04/Bergman-Roediger-1999_MemCog.pdf

Remembering

1932 · Frederic C. Bartlett

הספר שהציג את הזיכרון כשחזור מונחה סכמה, עם מתודולוגיה שלא הייתה עומדת בסטנדרט של היום ומסקנה ששרדה. לא אותר תרגום עברי.

The Seven Sins of Memory

2001 · Daniel L. Schacter

מיפוי שבעת אופני הכישלון של הזיכרון, כתוב בידי אחד החוקרים המרכזיים בתחום. לא אותר תרגום עברי.

Eyewitness Testimony

1979 · Elizabeth F. Loftus

הטקסט שהכניס את מחקר הזיכרון לאולם בית המשפט, ומכין את היחידה הבאה. לא אותר תרגום עברי.

זיכרונות שווא ועדי ראייה

1995 · מה שמחקר הזיכרון עשה לבית המשפט

ב 1995 פורסם ניסוי שבו הוצגו לעשרים וארבעה נבדקים אירועי ילדות: שלושה אמיתיים, שסופקו בידי בן משפחה מבוגר, ואחד בדוי — ללכת לאיבוד בקניון. כרבע מהם "נזכרו" באירוע שלא קרה. המספר הזה התגלגל לעולם כעובדה מוגמרת, וכדאי לדעת מה קרה לו מאז. ב־2017 אוחדו נתונייהם של שמונה מחקרי שתילה, 423 נבדקים בסך הכול, ו־30 אחוז סווגו כמי שפיתחו זיכרון שווא; בשכפול מוקדם־רישום מ־2023, על 123 נבדקים, השיעור היה 35 אחוז — כלומר גבוה מהמקור. אבל יש כאן גם ויכוח אמיתי: בהגדרה מחמירה יותר, שדורשת דימוי מנטלי ממשי ואמונה כנה של הנבדק, השיעורים יורדים ל־25 ואף ל־15 אחוז. ההבדל אינו טכני — הוא השאלה מה בכלל נחשב "זיכרון" ומה הוא סיפור שאדם בונה כדי לרצות את החוקר. וידוע גם שהשתילה עובדת הרבה פחות טוב ככל שהאירוע פחות סביר.

בשנות התשעים יצא הרעיון הזה מהמעבדה. טכניקות סוגסטיביות — היפנוזה, דימוי מודרך, פירוש חלומות — שימשו כדי "לשחזר" זיכרונות, ועל בסיסם הוגשו תביעות. הפער שנפתח אז בין מחקר לפרקטיקה לא נסגר: בסקרים שקובצו לאחרונה, כ־70 אחוז מהפסיכולוגים הקליניים מחזיקים באפשרות של זיכרון מודחק שאפשר לשחזר במדויק, ובקרב מי שנשאלו אחרי 2010 השיעור אף עלה ל־76 אחוז — לעומת 12.5 אחוז בלבד מבין חוקרי הזיכרון עצמם. בבתי המשפט נמצאו שיעורים דומים לאלה של הקלינאים בקרב מושבעים, שופטים ואנשי אכיפה. המונח עצמו, אגב, לא נעלם אלא החליף שם: במהדורה הנוכחית של מדריך האבחון האמריקני הוא מופיע אפס פעמים, בעוד ש"אמנזיה דיסוציאטיבית" מופיעה 75 פעמים.

הנתון שמכריע את הדיון בא ממקום אחר. מתוך זיכויי הדיאן־איי בארצות הברית, 69 אחוז — 252 מתוך 367 מקרים — כללו זיהוי שגוי של עד ראייה. זהו הגורם התורם הבולט ביותר, לפני שימוש שגוי במדע פורנזי והודאות שווא. הממוצע הוא ארבע עשרה שנות מאסר לכל מזוכה. ובדיעבד אותרו 165 מהעבריינים האמיתיים, שביצעו בינתיים 154 עבירות אלימות נוספות — כלומר המחיר של זיהוי שגוי משולם פעמיים.

מכאן באו רפורמות, וגם כאן הפקולטה מקבלת שיעור. הרפורמה המפורסמת ביותר — הצגת חשודים אחד־אחד במקום כולם יחד — נתמכה בשנות השמונים והתשעים ונשמעה מבוססת; ניתוח מאוחר יותר הראה שהיתרון לכאורה נובע מכך שהשיטה גורמת לעדים להיות **שמרניים** יותר ולא מבחינים טוב יותר, ובמחקר שדה לא נמצא הבדל. דוח האקדמיה הלאומית האמריקנית למדעים מ־2014 סירב במפורש להמליץ עליה. מה שכן אומץ: שהמעביר את המסדר לא יידע מי החשוד, שייאמר לעד במפורש שייתכן שהעברייני אינו במסדר, ושתתועד רמת הביטחון של העד **ברגע הזיהוי עצמו**.

והנקודה האחרונה היא תיקון לתיקון. במשך שנים לימדו שביטחון של עד אינו מנבא דיוק. ב־2017 הראתה סינתזה מחודשת שזה נכון רק לביטחון שנמדד באולם, חודשים אחרי, אחרי שהעד נחשף לאינספור רמזים.

בתנאים נקיים, ברגע הזיהוי: זיהוי בביטחון נמוך היה נכון ב־26.6 אחוז מהמקרים, וזיהוי בביטחון גבוה — ב־84.9 אחוז. אותו עד, אותו מסדר. מה שמתנה הוא מתי נשאלת השאלה.

למה זה ברשימה

כי זו הנקודה היחידה בפקולטה שבה מחקר פסיכולוגי שינה בפועל את חייהם של אנשים שיושבים בכלא.

שנה	1995
תחום	זיכרון
נושא	זיכרון ומשפט

להעמיק

אליזבת לופטוס: כמה אמין הזיכרון שלך

לצפייה

החוקרת עצמה מציגה את העבודה ואת ההשלכות המשפטיות והאתיות, עם תמליל בעברית.

https://www.ted.com/talks/elizabeth_loftus_how_reliable_is_your_memory/transcript?language=he

זיכרונות השווא ידברו

לקריאה

תיאור בעברית של ניסוי הקניון ושל המשמעות המשפטית של זיכרון מושגל.

<https://davidson.org.il/read-experience/sciencepanorama/זיכרונות-השווא-ידברו/>

ביטחון עד ראייה ודיוק הזיהוי: סינתזה חדשה

לקריאה

המאמר שמתקן את התיקון ומראה 84.9 אחוז דיוק בביטחון גבוה בתנאים נקיים. באנגלית.

<https://journals.sagepub.com/doi/full/10.1177/1529100616686966>

זיכויי די־אן־איי בארצות הברית

מקורות וארכיון

הנתונים הרשמיים: מספר הזיכויים, אחוז מקרי הזיהוי השגוי, שנות המאסר בממוצע ופילוח יתר הגורמים. באנגלית.

<https://innocenceproject.org/dna-exonerations-in-the-united-states/>

לזהות את האשם, דוח האקדמיה הלאומית 2014

מקורות וארכיון

הדוח שאימץ ניהול סמוי ותיעוד ביטחון מידי, וסירב להכריע בין מסדר סדרתי לבוזמני. באנגלית.

<https://www.innocenceproject.org/wp-content/uploads/2016/02/NAS-Report-ID.pdf>

Eyewitness Testimony

1979 · Elizabeth F. Loftus

הספר שהכניס את מחקר הזיכרון לבית המשפט וקבע את סדר היום למחקר שנעשה אחריו. לא אותר תרגום עברי.

Remembering Trauma

2003 · Richard J. McNally

הסקירה המאוזנת ביותר של מצב הראיות בשאלת זיכרון וטראומה, כתובה כהיסטוריה של מחלוקת מדעית. לא אותר תרגום עברי.

Convicting the Innocent

2011 · Brandon L. Garrett

ניתוח שיטתי של מאות זיכויי די־אן־איי ושל מה בדיוק השתבש בכל אחד מהם. לא אותר תרגום עברי.

איך באמת לומדים

2006 · שתי שיטות עובדות, והנפוצות אינן ביניהן

בניסוי מ-2006 קיבלו מאה ועשרים סטודנטים אותו טקסט. חלקם קראו אותו שוב, חלקם נבחנו עליו. אחרי חמש דקות הקוראים ניצחו: 81 אחוז מול 75. אחרי יומיים התהפכה התמונה, 54 מול 68, ואחרי שבוע — 42 מול 56. בניסוי שני, עם שלושה משטרי לימוד, מי שקרא ארבע פעמים זכר אחרי שבוע 40 אחוז, ומי שקרא פעם אחת ונבחן שלוש פעמים זכר 61.

ואז מגיע הממצא שהוא לב היחידה. את אותם נבדקים שאלו כמה הם צופים שיזכרו בעוד שבוע. מי שקרא ארבע פעמים נתן לעצמו את הציון הגבוה ביותר — והיה הגרוע ביותר. מי שנבחן נתן לעצמו את הציון הנמוך ביותר — וניצח. תחושת השליטה והשליטה בפועל לא רק שאינן זהות, הן מסודרות הפוך. זה מסביר למה כל כך קשה לשכנע אנשים להחליף שיטה: השיטה היעילה מרגישה גרועה בזמן אמת.

שני הדברים שעובדים ידועים ומדודים. אפקט הריווח — לפרוס את אותו זמן לימוד על פני ימים במקום לדחוס אותו — נבדק במטא-אנליזה של 184 מאמרים ו-14,811 נבדקים, והתוצאה על פני כל המרווחים היא 47.3 אחוז נכון מול 36.7. אותו חומר, אותו זמן, רק פרוס אחרת. יש בו גם עידון שכדאי להכיר: המרווח האופטימלי גדל ככל שרוצים לזכור לטווח ארוך יותר, ומי שרוצה לזכור חצי שנה צריך מרווחים של חודש ומעלה. האפקט השני הוא תרגול השליפה, כלומר לבחון את עצמך במקום לקרוא שוב: מטא-אנליזה של 61 מחקרים נותנת גודל אפקט של 0.50, והוא **גדל** ככל שמרווח הזמן עד המבחן גדול יותר. עם משוב הוא כמעט מכפיל את עצמו, 0.73 לעומת 0.39 בלעדיו.

ודווקא כאן שווה להיות הוגנים עם המספרים. מחקרים שפורסמו על אפקט הבחינה נתנו 0.58, ומחקרים שלא פורסמו נתנו 0.25 — פער אמיתי, שמצביע על הטיית פרסום. בתת-מדגם של מחקרים בעלי מינון גבוה שני הערכים מתקרבים ל-0.65, כלומר האפקט שורד את הבדיקה, בגודל צנוע יותר. ובמחקרים מקוונים עם נשירה של שישים עד שבעים אחוז הוא פשוט לא מופיע: הוא דורש מאמץ אמיתי, ולא סתם לחיצה על כפתור.

מול שתי השיטות האלה עומד אחד המיתוסים העיקשים בחינוך. הרעיון שלכל אדם יש סגנון למידה — חזותי, שמיעתי, תנועתי — ושהתאמת ההוראה לסגנון משפרת את הלמידה נקרא "השערת ההתאמה", ויש לו מבחן ברור: לחלק לומדים לפי סגנון, להקצות אותם באקראי לשיטות הוראה, ולראות אם השיטה הטובה שונה בין הקבוצות. בסקירה מ-2008 נמצא בכל הספרות **מחקר אחד** שהתקרב לתכנון הזה, וגם בו היו בעיות; המחקרים שכן תוכננו כראוי נתנו תוצאות סותרות להשערה. ובכל זאת, בסקירה שיטתית של 37 מחקרים על 15,405 מחנכים, 89 אחוז מאמינים בה — והאמונה חזקה **יותר** אצל סטודנטים להוראה מאשר אצל מורים ותיקים.

שימי לב להבחנה שלא כדאי לפספס: להעדפות יש קיום. אנשים באמת מעדיפים ללמוד כך או אחרת. מה שאין לו ראיות הוא שהתאמת ההוראה להעדפה מועילה. וגם הסימון בטוש וקריאה חוזרת לא הוכחו כמזיקים — הם דורגו כבעלי תועלת נמוכה, כי הם קשים ליישום נכון והתוצאות מהן לא עקביות. הבעיה אינה שהן רעות. הבעיה היא שהן הנפוצות ביותר, ושתי המועילות אינן.

למה זה ברשימה

כי התחושה שאת לומדת טוב והלמידה בפועל נעות בכיוונים הפוכים, וזה נמדד בניסוי אחד נקי.

שנה	2006
תחום	זיכרון
נושא	למידה ותרגול

להעמיק

בחינה מחודשת

לקריאה

סקירה בעברית של אפקט הבחינה ושל הדירוג בין עשר טכניקות הלימוד הנפוצות.

<https://davidson.org.il/read-experience/scientificamerican/בחינה-מחודשת/>

למידה והוראה אינן נשענות על סגנונות למידה

לקריאה

הסבר עברי קצר וברור לכך שהשערת ההתאמה נשללה, כתוב בשפה פשוטה במיוחד.

<https://kids.frontiersin.org/he/articles/10.3389/frym.2020.00110-he>

סגנונות למידה: מושגים וראיות

לקריאה

המאמר שהגדיר איזה תכנון ניסויי דרוש כדי לבחון את ההשערה, ומצא מחקר אחד בלבד שכמעט עמד בו. באנגלית.

<https://journals.sagepub.com/doi/10.1111/j.1539-6053.2009.01038.x>

עד כמה נפוצה האמונה בסגנונות למידה

לקריאה

סקירה שיטתית של 37 מחקרים ו-15,405 מחנכים, ללא סימן לירידה באמונה לאורך העשור. באנגלית.

<https://www.frontiersin.org/journals/education/articles/10.3389/feduc.2020.602451/full>

למידה מוגברת בחינה, 2006

מקורות וארכיון

הניסוי שבו הקריאה החוזרת מנצחת אחרי חמש דקות ומפסידה אחרי שבוע, כולל התחזיות ההפוכות של הנבדקים. באנגלית.

https://colinallen.dnsalias.org/Readings/2006_Roediger_Karpicke_PsychSci.pdf

Make It Stick

2014 · Peter Brown, Henry Roediger & Mark McDaniel

כתוב בידי שניים מהחוקרים שערכו את הניסויים ביחידה, ומיועד ללומד ולא לחוקר. לא אותר תרגום עברי.

?Why Don't Students Like School

2009 · Daniel T. Willingham

עוסק ישירות בפער בין תחושת הלמידה ללמידה בפועל, מאת אחד ממחברי הדירוג. לא אותר תרגום עברי.

How We Learn

2014 · Benedict Carey

עיתונות מדע קריאה על ריווח, שליפה והשהיה, לקורא שרוצה את זה בסיפור ולא בטבלה. לא אותר תרגום עברי.

שתי המערכות

1974 · כהנמן, טברסקי, ומה שרד מההטיות

ב1974 פרסמו עמוס טברסקי ודניאל כהנמן מאמר בן שמונה עמודים שטען שבני אדם אינם מחשבים הסתברויות אלא משתמשים בקיצורי דרך — היוריסטיקות — שעובדים לרוב ונכשלים באופן שיטתי. שלוש מהן קיבלו שם: ייצוגיות, זמינות ועיגון. הדגמת העיגון היא עדיין ההדגמה היפה ביותר במדעי ההתנהגות: סובבו גלגל מזל מזויף, שאלו נבדקים אם אחוז מדינות אפריקה באו"ם גדול או קטן מהמספר שיצא, ואז ביקשו אומדן. מי שקיבל 10 אמד בחציון 25; מי שקיבל 65 אמד 45. אותה שאלה, אותם אנשים, עשרים נקודות הפרש.

טברסקי מת מסרטן ב־1996, בן חמישים ותשע. פרס נובל לכלכלה הוענק ב־2002 לכהנמן — פרס נובל אינו ניתן לאחר המוות — והודעת הוועדה נקבה בשמו של טברסקי במפורש.

וכאן מגיע החלק שהפקולטה הזאת עוסקת בו. פרויקט Many Labs בדק את העיגון בשלושים ושישה מדגמים ו־6,344 נבדקים, והאפקט לא רק שוחזר אלא יצא **גדול יותר** מהמקור — גדלי אפקט של 2.05 עד 2.60 במקום 0.93. באותו פרויקט עצמו נפלו לגמרי שני ממצאים אחרים, שניהם מתחום הפריימינג החברתי: חשיפה לדגל אמריקני שאמורה להזיז עמדות פוליטיות ירדה מגודל אפקט 0.50 לאפס נקודה אחת, וחשיפה לכסף ירדה מ־0.80 למינוס אפס נקודה אחת. עשרה מתוך שלושה עשר האפקטים שנבדקו עמדו יפה. הכישלונות אינם מפוזרים בשווה — הם מרוכזים במקום אחד.

המקום הזה הוא בדיוק הפרק הרביעי של "לחשוב מהר לחשוב לאט", ספרו של כהנמן מ־2011, שעוסק בפריימינג חברתי. ב־2017 פרסם חוקר בשם אולריך שימאק ניתוח של שנים עשר המאמרים ושלושים ואחת העבודות שכהנמן ציטט באותו פרק: מאה אחוז מהתוצאות היו מובהקות, אבל העוצמה הסטטיסטית החציונית הייתה 57 אחוז — כלומר לפי חוקי ההסתברות היו אמורות להיכשל כמה מהן. שלמות כזאת היא סימן, ולא סימן טוב. כהנמן, אז בן שמונים ושתיים, הגיב בתגובה פומבית בבלוג. שלוש שורות ממנה שוות ציטוט: "אני מקבל את המסקנות הבסיסיות של הבלוג הזה"; "מה שהבלוג צודק בו לחלוטין הוא שהאמנתי יותר מדי במחקרים חלשי עוצמה"; ו"שיניתי את דעתי על גודלם של אפקטי הפריימינג ההתנהגותיים — הם אינם יכולים להיות גדולים וחסינים כפי שהפרק שלי הציג". הוא גם אמר במפורש שלא חדל להאמין לאף אחד מהמחקרים עצמם. זו לא נסיגה; זו הודאה מדויקת.

כדאי לדעת שכהנמן ראה את זה מגיע. ב־2012 שלח מכתב פתוח לחוקרי הפריימינג ובו המשפט "אני רואה תאונת רכבת מתקרבת. אני צופה שהקורבנות הראשונים יהיו צעירים שמחפשים עבודה", והציע שרשרת שחזורים בין מעבדות. חמש שנים אחר כך הוא כתב שהתאונה קרתה.

ולבסוף, המסגרת שכולם מכירים. "מערכת 1" ו"מערכת 2" אינן המצאה של כהנמן — הוא מייחס אותן במפורש לסטנוביץ' ולווסט — והוא כותב בספר עצמו, במילים שלו, שהן "דמויות בדיוניות", שאינן מערכות במובן

המקובל, ושאינן אזור במוח שאפשר לקרוא לו בית של אחת מהן. ב־2018 פורסמה ביקורת שיטתית שטוענת שהחלוקה הדו־מערכתית "נטולת תמיכה אמפירית, סותרת ממצאים מבוססים ואינה עקבית בתוך עצמה". השורה התחתונה ההוגנת: השיפוט וההחלטות שרדו יפה. הפריימינג החברתי לא. והמטפורה שמוכרת את שניהם היא מטפורה.

למה זה ברשימה

כי זה התחום שממנו צמחה הכלכלה ההתנהגותית — וגם המקום שבו הפסיכולוג המפורסם בעולם הודה שטעה.

שנה	1974
תחום	חשיבה והחלטות
נושא	היוריסטיקות והטיות

להעמיק

המאמר המקורי, 1974

לקריאה

שמונה עמודים שבהם מוצגות שלוש ההיוריסטיקות וניסוי גלגל המזל, בשפה פשוטה להפתיע. באנגלית.

<https://bear.warrington.ufl.edu/brenner/mar7588/Papers/tversky-kahneman-science-1974.pdf>

הפסיכולוג ששינה את המחשבה הכלכלית

לקריאה

פורטרט בעברית של עמוס טברסקי, כולל נסיבות מותו והקשר לפרס נובל שהוענק שש שנים אחריו.

<https://davidson.org.il/read-experience/sciencehistory/הפסיכולוג-ששינה-את-המחשבה-הכלכלית/>

שחזור של תאונת רכבת

לקריאה

הניתוח של הפרק הרביעי — ובתגובות שמתחת, דבריו של כהנמן עצמו, מילה במילה. באנגלית.

<https://replicationindex.com/2017/02/02/reconstruction-of-a-train-wreck-how-priming-research-went-of-the-rails/>

"האמנתי יותר מדי במחקרים חלשי עוצמה"

לקריאה

הדיווח העיתונאי על ההודאה, עם ההקשר ועם הציטוטים המלאים. באנגלית.

<https://retractionwatch.com/2017/02/20/placed-much-faith-underpowered-studies-nobel-prize-winner-admits-mistakes>

Many Labs 1, המאמר המלא

מקורות וארכיון

שלושה עשר ממצאים בשלושים ושישה מדגמים: מה גדל, מה נשאר, ומה ירד לאפס. באנגלית.

<https://stanford.edu/~knutson/jdm/klein14.pdf>

לחשוב מהר, לחשוב לאט

דניאל כהנמן · 2011

הספר שהביא את כל התחום לקהל הרחב, ויצא בעברית בעריכה מדעית של מיה בריהלל. כדאי לקרוא אותו עם היחידה הזאת ביד — ובמיוחד את הפרק הרביעי.

Judgment under Uncertainty

1982 · Kahneman, Slovic & Tversky

קובץ המאמרים שהמאמר מ־1974 פותח אותו, ומראה כמה רחב היה הפרויקט כבר אז. לא אותר תרגום עברי.

The Undoing Project

2016 · Michael Lewis

סיפור הידידות בין כהנמן לטברסקי, שהוא גם הסבר טוב לאיך נראתה עבודה משותפת כזאת. לא אותר תרגום עברי.

מסגור ושנאת הפסד

1979 · אחד הממצאים שהחזיקו, ואחד שמתווכחים עליו עכשיו

הניסוי המפורסם נקרא "מחלת אסיה". מציגים לנבדקים מחלה שצפויה להרוג שש מאות איש, ומציעים שתי תוכניות. בניסוח הראשון: תוכנית א' מצילה מאתיים; תוכנית ב' נותנת שליש סיכוי להציל את כולם ושני שלישים סיכוי להציל אפס. שבעים ושניים אחוז בוחרים בוודאות. בניסוח השני: תוכנית ג' — ארבע מאות ימותו; תוכנית ד' — שליש סיכוי שאיש לא ימות. שבעים ושמונה אחוז בוחרים בהימור. התוצאות זהות לחלוטין. רק המילים שונות.

הממצא הזה הוא אחד הניצחונות הגדולים של הפסיכולוגיה בבדיקה מחדש. בפרויקט Many Labs הוא שוחזר בשלושים ושישה מדגמים, וגודל האפקט ירד מ 1.13 ל 0.62 — בערך לחצי, אבל חד-משמעי. בפרויקט ההמשך, על 7,228 נבדקים, יחס הסיכויים ירד מ 4.96 ל 2.06 , עם מובהקות אסטרונומית. כלומר: האפקט אמיתי, והוא קטן בערך בחצי ממה שדווח במקור. שני המשפטים האלה נכונים יחד, וזה בדיוק הדפוס שחוזר בכל הפקולטה.

התיאוריה שמאחורי זה, תורת הערך, פורסמה ב־1979 ועליה ניתן פרס נובל. שווה לדעת על מה היא נבנתה: שבעים ושניים סטודנטים וחברי סגל ישראלים שענו על שאלות היפותטיות על דף. בלי כסף אמיתי. ובמאמר ההוא אין שום מספר לשנאת ההפסד — יש משפט איכותי, שהעוגמה שבאובדן סכום גדולה מההנאה שבזכייה באותו סכום. המספר המפורסם, שנאת ההפסד היא בערך פי שניים, הגיע רק ב־1992, מחציון של מחקר אחד: 2.25.

והנה מה שקרה לו. במטא־אנליזה מ־2024 שאיחדה תשעה עשר מאגרי נתונים ו־1,192 נבדקים, האומדן המשוקלל הוא 1.31, ובשנים עשר מתוך תשעה עשר המאגרים רווח הסמך כלל את האפשרות שאין שנאת הפסד כלל. השונות בין המחקרים עצומה, והמחברים כותבים בפירוש שחלק ניכר מהנתונים באיכות ירודה. ממחקר אחר, ישראלי, עולה תלות בגודל הסכום: בסיכון של שישה דולר אין שנאת הפסד בכלל, בארבעים דולר היא 1.16, ובמאה דולר 1.54. אנשים אינם שונאים הפסד; הם שונאים הפסד גדול.

ואז הגיע הערעור העקרוני. ב־2018 טענו גל ורוקר שהראיות אינן תומכות בכך שהפסדים משפיעים יותר מרווחים "בסך הכול", ושחלק ניכר ממה שמיוחס לשנאת הפסד הוא בעצם נטייה לאי־פעולה. ההדגמה החדה שלהם: יותר משמונים וחמישה אחוז מהנבדקים סירבו להחליף שטר של עשרים דולר בשטר אחר של עשרים דולר — מצב שבו אין שום הפסד אפשרי. התשובה מבפנים לא איחרה: מחקר על 17,720 נבדקים, שכותרתו "הידיעות על מותה הוגזמו", מצא שידע, ניסיון והשכלה מחלישים את שנאת ההפסד — "אבל אנשים בכל רמות הידע, הניסיון וההשכלה היו שונאי הפסד".

דוגמה אחרונה שמראה כמה זה עדין. אפקט ההקצאה — שברגע שדבר שייך לך הוא שווה בעיניך יותר — הודגם ב־1990 על ארבעים וארבעה סטודנטים וספל בשש דולר: המוכרים רצו 5.25 בחציון, הקונים הציעו 2.75.

ובמקום אחת עשרה עסקאות צפויות התרחשו בערך שתיים. ב-2005 חזרו על הניסוי עם ארבע בקורות בו־זמנית — מנגנון תמחור שמעודד גילוי אמת, הסבר מפורט עם דוגמאות, ארבעה עשר סבבי תרגול משולמים, ואנונימיות — והפער נעלם לחלוטין בכל שלוש הגרסאות. זה לא סוגר את הוויכוח, כי יש טענות לכישלון שחזור גם של המחקר הזה. זה כן אומר שהתשובה הכנה כאן היא "תלוי", וששנאת הפסד אינה חוק טבע אלא תופעה עם מודרטורים.

למה זה ברשימה

כי שני הרעיונות האלה נסעו יחד ארבעים שנה, ורק כשבדקו אותם בנפרד התברר שאחד מהם חזק הרבה יותר מהשני.

שנה	1979
תחום	חשיבה והחלטות
נושא	תורת הערך

להעמיק

אובדן שנאת ההפסד, 2018

לקריאה

הטיעון המלא נגד שנאת הפסד כעיקרון אוניברסלי, כולל ניסוי החלפת שטר העשרים. באנגלית.

<https://statmodeling.stat.columbia.edu/wp-content/uploads/2018/06/Loss-of-Loss-Aversion.pdf>

הידיעות על מותה הוגזמו, 2020

לקריאה

התשובה האמפירית, על 17,720 נבדקים: שנאת ההפסד קיימת בכל הקבוצות, ויש לה מודרטורים ברורים. באנגלית.

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3503581

מטא־אנליזה של שנאת הפסד, 2024

לקריאה

כאן נקבע שהמקדם האמיתי קרוב ל-1.31 ולא ל-2.25, וששבשני שלישים מהמאגרים אי אפשר לשלול אדישות. באנגלית.

<https://www.sciencedirect.com/science/article/pii/S0167487024000485>

על כשלי החשיבה האנושית

לקריאה

מקור עברי נגיש על שנאת הפסד ותורת הערך; שימי לב שאינו נוגע כלל במחלוקת שהיחידה מציגה.

<https://www.ynet.co.il/economy/article/r1ztdafyc>

המאמר המקורי על מסגור, 1981

מקורות וארכיון

"מחלת אסיה" בניסוח המדויק שלה, עם מספרי הנבדקים והאחוזים בכל אחד משני הניסוחים. באנגלית.

<https://psych.hanover.edu/classes/cognition/papers/tversky81.pdf>

ניסוי הספלים, 1990

מקורות וארכיון

אפקט ההקצאה בגרסתו הקלאסית, כולל בקרת האסימונים שהיא החלק המשכנע באמת. באנגלית.

https://web.mit.edu/curhan/www/docs/Articles/15341_Readings/Behavioral_Decision_Theory/Kahneman_et_al_1990_Experimental_tests.pdf

ספרים

לחשוב מהר, לחשוב לאט

דניאל כהנמן · 2011

החלק הרביעי בספר הוא תורת הערך והמסגור, בניסוח של מי שכתב את המאמר המקורי. בעברית, בעריכה מדעית של מיה בריהלל.

Choices, Values, and Frames

Kahneman & Tversky (עורכים) · 2000

קובץ שמאגד את מאמרי 1979 ו-1981 יחד עם התיקונים שהמחברים עצמם הכניסו אחר כך. לא אותר תרגום עברי.

Misbehaving

Richard H. Thaler · 2015

ההיסטוריה של הכלכלה ההתנהגותית מבפנים, מאת מי שהמציא את הניסוי עם הספלים. לא אותר תרגום עברי.

ביטחון יתר וניבוי

2005 · הממצא שמחזיק שבעים שנה ואיש אינו מקשיב לו

ב1954 פרסם פול מיל ספר קטן ומטריד שבו השווה את שיפוטו של קלינאי מאומן לנוסחה סטטיסטית פשוטה. הנוסחה ניצחה. שלושים ושתיים שנה אחר כך כתב מיל את המשפט שמסכם את הקריירה שלו: "אין מחלוקת במדעי החברה שבה גוף כה גדול של מחקרים מגוונים יוצא כה אחד לאותו כיוון". באותה נקודה היו כתשעים מחקרים, מניבוי תוצאות פוטבול ועד אבחון מחלות כבד, ובקושי חצי תריסר מהם תמכו ולו במעט בשיפוט הקליני.

ב־2000 נעשתה מטא־אנליזה של מאה שלושים ושישה מחקרים. הניקוד: בארבעים ושבעה אחוז מהמחקרים הנוסחה עדיפה במידה ניכרת, בארבעים ושבעה אחוז השניים שקולים בערך, ובשישה אחוז — שמונה מחקרים — הקלינאי מנצח. בשבעה מתוך אותם שמונה, לקלינאי היה **יותר** מידע מאשר לנוסחה. היתרון הממוצע של הנוסחה קטן, כעשרה אחוזים, אבל הוא עקבי להחריד; והמקרה היחיד שבו השיפוט האנושי נחלש במיוחד הוא כשהוא מבוסס על ראיון אישי. מיל עצמו מנה את הסיבות שבגללן התחום מתעלם: בורות סטטיסטית, פחד מייתור מקצועי, איום על התפיסה העצמית, תחושה שהשיטה מנטרלת את האנושי — ופחד ממחשבים.

הצד השני של המטבע נבדק במחקר אורך של עשרים שנה. פיליפ טטלוק אסף 82,361 תחזיות מ־284 מומחים שהתפרנסו מפרשנות על מגמות פוליטיות וכלכליות. המומחים בקושי היכו משכילים לא מומחים, ואיש מהם לא היכה כללי חזוי פשוטים; בשאלות עם שלוש אפשרויות הם לעיתים קרובות הצליחו פחות מאשר מי שהיה מניח שכל אפשרות שווה. מדד הכיול מטריד במיוחד: כשמומחים אמרו "שמונים וחמישה אחוז", הדבר קרה בכ־65 עד 70 אחוז מהמחקרים. חמישה עשר אחוז ממה שהכריזו כבלתי אפשרי קרה; יותר מרבע ממה שהכריזו כוודאי לא קרה. וטטלוק גם הבחין בין "קיפודים", שמחזיקים ברעיון גדול אחד, ל"שועלים", שמלקטים הרבה שיקולים קטנים ומשנים דעה — השועלים ניצחו.

וכאן חשוב לעצור, כי המשפט הפופולרי שנדבק לעבודה הזאת — שמומחים אינם טובים משימפנזה שזורק חצים — הוא עיוות, וטטלוק עצמו מכנה אותו שטות: "מה שהמחקר שלי הראה הוא שהמומחה הממוצע הצליח מעט יותר מניחוש **בשאלות רבות** שהצבתי. רבות אינו כל". ההמשך שלו מוכיח זאת. בתחרות חזוי שערך ארגון מודיעין אמריקני בין 2011 ל־2015, עם כחמש מאות שאלות על ענייני עולם, גייס טטלוק יותר מעשרים אלף חובבים. הקבוצה שלו היכתה את קבוצת הביקורת בשישים אחוז בשנה הראשונה ובשבעים ושמונה בשנייה, וגברה על צוותי אוניברסיטאות מתחרים בשלושים עד שבעים אחוז — וגם על אנליסטים מקצועיים בעלי גישה לחומר מסווג. שלושה דברים ייצרו את זה: אימון קצר בהסתברות, עבודה בצוותים, וקידום שני האחוזים הטובים ביותר לקבוצות עילית. ומעל הכול — הם שמרו על הדיוק שנתיים ברציפות, בניגוד לציפייה שהכול יחזור לממוצע.

שתי הערות אחרונות, שתיהן חשובות. הראשונה: "ביטחון יתר" אינו דבר אחד. בעבודה מ־2008 הופרד המושג לשלושה — הערכת יתר של הביצוע שלך, הצבה עצמית מעל למוצא, ודיוק יתר, כלומר רווחי ודאות צרים מדי. השניים הראשונים נעו בניסוי בכיוונים **הפוכים**, במתאם של מינוס 0.64: במשימות קלות אנשים מעריכים את עצמם בחסר אך מציבים את עצמם מעל האחרים, ובמשימות קשות ההפך. השנייה: הטיית בדיעבד, התחושה ש"ידעתי שזה יקרה", נמדדה במטא־אנליזה של 122 מחקרים וגודלה קטן, 0.17 — והמחברים עצמם כתבו שהיא "אולי אינה מדאיגה כפי שמקובל לחשוב". גם כאן, המספר האמיתי צנוע יותר מהמוניטין.

למה זה ברשימה

כי כאן הכיוון הפוך מכל שאר הפקולטה: לא ממצא שקרס, אלא ממצא שמסרב לקרוס ושהמקצוע מתעלם ממנו בעקביות.

שנה	2005
תחום	חשיבה והחלטות
נושא	שיפוט מומחים

להעמיק

שיפוט פוליטי של מומחים

לקריאה

תיאור המחקר בן עשרים השנה ואבחנת השועל והקיפוד, מעמוד ההוצאה לאור. באנגלית.

<https://press.princeton.edu/books/hardcover/9780691178288/expert-political-judgment>

משבר המומחים

לקריאה

מקור עברי שמביא את מספריו של טטלוק ואת פרויקט השיפוט הטוב, כולל נתוני הכיול.

<https://alaxon.co.il/article/המומחים>

ניבוי קליני מול ניבוי מכני, 2000

מקורות וארכיון

המטא־אנליזה של 136 מחקרים, עם הניקוד המלא ועם הממצא שהראיון האישי דווקא מחליש את השיפוט. באנגלית.

<http://zaldlab.psy.vanderbilt.edu/resources/wmg00pa.pdf>

פול מיל: הסיבות והתוצאות של ספרי הקטן והמטריד

מקורות וארכיון

מיל עצמו, שלושים ושתיים שנה אחרי, מסביר למה איש אינו מקשיב — כולל פחד ממחשבים. באנגלית.

<https://meehl.umn.edu/sites/meehl.umn.edu/files/files/132causeseffects.pdf>

הצרה עם ביטחון היתר, 2008

מקורות וארכיון

שלושה סוגים שאינם אותו דבר, ולפעמים מושכים לכיוונים מנוגדים. באנגלית.

https://healy.econ.ohio-state.edu/papers/Moore_Healy-TroubleWithOverconfidence.pdf

ספרים

Superforecasting

2015 · Philip Tetlock & Dan Gardner

הספר שמסביר מה בדיוק עשו חזאי העל, ושבט טטלוק מתקן במפורש את משפט השימפנזה המיוחס לו. לא אותר תרגום עברי.

Expert Political Judgment

2005 · Philip Tetlock

המחקר עצמו: עשרים שנה, 284 מומחים, 82 אלף תחזיות. לא אותר תרגום עברי.

רעש

דניאל כהנמן, אוליבייה סיבוני וקאס סאנסטיין · 2021

על השונות בשיפוט של מומחים באותו תפקיד ועל מה שעושים איתה, בעברית ובעריכה מדעית של מיה בריהלל.

מבחן המרשם

2018 · המתאם לא נעלם. הוא התכווץ ושינה משמעות

הסיפור שכולם מכירים: מושיבים ילד בן ארבע מול ממתק, אומרים לו שאם יחכה עד שהמבוגר יחזור יקבל שניים, ויוצאים. מי שמחכה — יצליח בחיים. זה נמדד, ויש מתאמים, ויש ספר.

נתחיל במה שקרה במעבדה בפועל, כי גם זה כבר לא מדויק. הניסויים המקוריים בגן הילדים של סטנפורד, בסוף שנות השישים ובתחילת שנות השבעים, לא היו על ניבוי כלל. הם היו על **אסטרטגיית קשב** האם הממתק גלוי או מכוסה, ומה קורה כשמלמדים ילד לדמיין שהוא תמונה במסגרת. ילדים שקיבלו את ההנחיה הזאת חיכו כמעט שמונה עשרה דקות, ואחד מהם הסביר: "אי אפשר לאכול תמונה". גם הממתק לא תמיד היה מרשם — בניסוי הראשון, על שלושים ושניים ילדים, היו עוגיות חיות ומקלות מלוחים, ולעיתים קרובות הילד בחר בעצמו.

טענת הניבוי המפורסמת הגיעה ב־1990, וכדאי לדעת על מה היא נשענה. מתוך כשש מאות וחמישים ילדים שעברו בגן, נשלחו שאלונים ל־404 משפחות, 185 השיבו, ורק 94 מסרו ציוני מבחן כניסה לקולג'. המתאמים המפורסמים, 0.57 ו־0.42, מגיעים מתת־קבוצה של כשלושים וחמישה עד ארבעים ושמונה ילדים.

ב־2018 נעשה שחזור מושגי על מדגם גדול ומגוון בהרבה: 918 ילדים, ומתוכם תת־מדגם מוגדר מראש של 552 ילדים לאימהות ללא תואר. התוצאה אינה "לא נמצא כלום". נמצא מתאם, באותו כיוון, ובערך חצי מגודלו במקור. ואז נוספו בקורות: רקע משפחתי וסביבת הבית הורידו את המקדם מ־0.24 ל־0.10, ותוספת של יכולת קוגניטיבית והתנהגות שנמדדו באותו גיל הורידה אותו ל־0.05 — לא מובהק. ועוד ממצא ששווה יותר מכל השאר: כמעט כל האות הניבויי יושב בשאלה אם הילד החזיק **עשרים שניות***, לא חמש עשרה דקות. התמונה של הילד המתייסר בדקה השמינית אינה עושה שום עבודה סטטיסטית.

ההמשך חד עוד יותר. במעקב עד גיל עשרים ושש, על 702 נבדקים, כמעט כל המתאמים הגולמיים נעלמו עם בקורות; מה ששרד באופן עקבי היה מדד מסת גוף נמוך יותר. ובמחקר על הקוהורט המקורי עד גיל ארבעים ושש, על 113 איש ואחת עשרה תוצאות כלכליות, המתאם הממוצע של מדד ההמתנה בגן היה 0.02. אחד מהחותמים על המחקר ההוא היה וולטר מישל עצמו, שהמציא את הניסוי; הוא נפטר בזמן עריכת המאמר.

מישל, אגב, לא הגן על הקריאה הפופולרית — הוא תקף אותה. בראיון לפני מותו: "הסיפור האיקוני הזה הפוך לגמרי — שהעתיד שלך נמצא במרשם — כי הוא לא".

ומה כן מסביר את ההמתנה? ב־2013 נערך ניסוי קטן ואכזרי ביופיו: לפני המרשם הבטיח המבוגר לילד שני דברים במהלך משימת ציור. בקבוצה אחת הוא קיים; בשנייה הוא הפר את שתי ההבטחות ואמר "טעיתי", אין לנו מדבקות אחרות". בקבוצת ההבטחות שנשמרו חיכו תשעה מתוך ארבעה עשר את כל רבע השעה; בקבוצה

השנייה — אחד מתוך ארבעה עשר, ומשך ההמתנה הממוצע ירד משתים עשרה דקות לשלוש. ההמתנה אינה תכונת אופי. היא החלטה רציונלית לגבי השאלה אם אפשר לסמוך על מי שהבטיח.

אז לא, המבחן לא "הופרך". אל תכתבי ואל תספרי את זה כך. המתאם קיים, קטן, ומוסבר כמעט כולו על ידי דברים שהילד הביא איתו לחדר. מה שקרס הוא הפרשנות: שכוח הרצון הוא כמות קבועה, שהוא קובע את העתיד, ושאימון שלו יזיז תוצאות חיים.

למה זה ברשימה

כי זה המקרה שמלמד את הדבר הקשה יותר: לא ממצא שקרס, אלא ממצא שנשאר נכון וההסבר שלו התהפך.

שנה	2018
תחום	שליטה עצמית
נושא	דחיית סיפוקים

להעמיק

הילדים שלמדו להתאפק

לקריאה

סקירה בעברית של המבחן, של מה שהשתנה בין הדורות, ושל הביקורת החדשה.

<https://davidson.org.il/read-experience/sciencenews/הילדים-שלמדו-להתאפק/>

מרשמלו מועד גימל

לקריאה

מסה בעברית שהולכת אחרי מישל ואחרי הקריסה של הפרשנות, כולל הקריאה החברתית-כלכלית.

https://alaxon.co.il/?post_type=article&p=145692

נסו להתאפק מלפרש לא נכון את מבחן המרשמלו

לקריאה

הצד שמזהיר מהתיקון היתר: המתאם נשאר באותו כיוון ומובהק, והתקשורת המציאה הפרכה שלא הייתה. באנגלית.

<https://behavioralscientist.org/try-to-resist-misinterpreting-the-marshmallow-test/>

הספר לוולטר מישל

לקריאה

כאן נמצא הציטוט שבו ממצא הניסוי אומר שהסיפור שנבנה עליו הפוך לגמרי. באנגלית.

<https://www.npr.org/sections/health-shots/2018/09/21/650015068/remembrance-for-walter-mischel-psychologist-who-devised-the-marshmallow-test>

ביקור חוזר במבחן המרשמלו, 2018

מקורות וארכיון

השחזור המלא, עם הטבלאות שמראות איך המקדם צונח מ-0.24 ל-0.05 כשמוסיפים בקורות. באנגלית.

<https://scispace.com/pdf/revisiting-the-marshmallow-test-a-conceptual-replication-1hgo0biknc.pdf>

ניסוי ההבטחה שהופרה, 2013

מקורות וארכיון

הניסוי שבו הפרת שתי הבטחות מורידה את ההמתנה משתים עשרה דקות לשלוש. באנגלית.

<https://www.strategian.com/fulltext/Kidd2013.pdf>

ספרים

The Marshmallow Test

2014 · Walter Mischel

הספר שבו מישל עצמו הציג את העבודה לקהל הרחב — ארבע שנים לפני השחזור שהפך את הפרשנות שבו. לא אותר תרגום עברי.

Scarcity

2013 · Sendhil Mullainathan & Eldar Shafir

המסגרת התיאורטית לקריאה החברתית-כלכלית: מה מחסור עושה לקשב ולהחלטות. לא אותר תרגום עברי.

Science Fictions

2020 · Stuart Ritchie

התמונה הרחבה של משבר השחזור, שהיחידה הזאת והבאה אחריה הן שתי דוגמאות שונות מאוד שלו. לא אותר תרגום עברי.

דלדול האגו

2016 · איך נראית קריסה אמיתית של ממצא

הניסוי מ-1998 בנוי כמו הצגה. אופים עוגיות שוקולד בחדר, כדי שהריח ימלא אותו. על השולחן מניחים שני מגשים: עוגיות חמות וממתקי שוקולד מצד אחד, קערת צנוניות מהצד השני. שישים ושבעה סטודנטים שהתבקשו לצום מגיעים, ומחולקים לשלוש קבוצות — לאכול רק צנוניות, לאכול שוקולד חופשי, או כלום. אחר כך כולם מקבלים חידה גיאומטרית שאין לה פתרון, ומודדים כמה זמן הם מתעקשים.

קבוצת הצנוניות החזיקה 8.35 דקות. קבוצת השוקולד — 18.9. קבוצת הביקורת — 20.86. פער של יותר מעשר דקות, מכך שאדם התאפק מלאכול עוגיה. מכאן נולד רעיון שכבש את העולם: כוח הרצון הוא משאב מוגבל, כמו שריר שמתעייף או מכל שמתרוקן, ושימוש בו במקום אחד מחליש אותו במקום אחר. ארבעת הניסויים באותו מאמר כללו יחד מאתיים ועשרים איש. עליהם נבנו יותר משש מאות מחקרים ושני רבי מכר.

הגרסה שהגיעה לציבור הלכה צעד נוסף: שהמכל מתמלא בגלוקוז, ושמשיקה ממותק מחזיר כוח רצון. ב-2010 סיכמה מטא-אנליזה שמונים ושלושה מחקרים וקבעה גודל אפקט 0.62, בינוני עד גדול, וגם אישרה לכאורה את סיפור הסוכר.

וכאן מתחיל התיקון, בשלושה שלבים בלתי תלויים. ראשית, ב-2014, ניתוח הטיית פרסום על אותו גוף מחקרים: כשמתקנים לפי הקשר בין גודל המדגם לגודל האפקט, האומדן מתכווץ עד לאפס ואף מתחתיו. המחברים כתבו שהאפקט "אינו ניתן להבחנה מאפס" ושהראיות "הוערכו ביתר באופן חמור". שנית, ב-2016, נבדק סיפור הגלוקוז בנפרד: שלושים ושמונה השוואות מתוך תשעה עשר מאמרים, והתפלגות התוצאות התגלתה שטוחה — כלומר, לפי הכלי הסטטיסטי הזה, למחקרים האלה "אין כמעט ערך ראייתי".

ושלישית — הבדיקה הישירה. עשרים ושלוש מעבדות, 2,141 נבדקים, פרוטוקול שנקבע מראש ונרשם לפני האיסוף. התוצאה: גודל אפקט 0.04, עם רווח סמך שכולל אפס. שווה לשים לב לשני דברים. האחד: זה בדיוק מה שניתוח הטיית הפרסום ניבא שנתיים קודם. השני: מתוך עשרים ושלוש המעבדות, עשרים ושתיים צפו לפני ההרצה שהאפקט **כן** ישוחרר. רבות מהן היו מעבדות מומחים בשליטה עצמית. האינטואיציה המקצועית של התחום כולו לא תפסה את המציאות.

המחברים המקוריים לא קיבלו זאת. טענתם: המשימה שנבחרה לשחזור — מחיקת אותיות לפי כלל — לא מייצרת הרגל שצריך להתגבר עליו, ולכן לא באמת מדללת דבר. אבל הם גם כתבו משפט שראוי לצטט: "בדיעבד, ההחלטה להשתמש בהליכים חדשים ולא בדוקים לפרויקט שחזור גדול הייתה טיפשית... חשבנו שחוסנו של דלדול האגו יתגבר על כל חולשה בשיטה החדשה. זו הייתה טעות מצערת". ב-2021 נערכה בדיקה נוספת שעונה בדיוק להתנגדות הזאת, עם שתי משימות אחרות: שלושים ושש מעבדות, 3,531 נבדקים. גודל האפקט המאושר מראש: 0.06, לא מובהק. הנתונים היו פי ארבעה סבירים יותר תחת ההנחה שאין אפקט.

הוויכוח לא הסתיים בנימוס. באומייסטר ממשיך לטעון עד היום שדלדול האגו הוא מהממצאים החזקים בפסיכולוגיה החברתית, וחוקרים אחרים מכנים את הטענה הזאת מנותקת מהראיות. מייקל אינזליכט, שבנה את הקריירה שלו על התיאוריה וקיבל עליה פרס, כתב בדיעבד: "שוב ושוב ניסיתי. שוב ושוב נכשלתי", ו"הרגשתי כמו רמאי".

ושורה אחרונה של הגינות: מה שקרס הוא המודל הספציפי — מכל שמתרוקן תוך דקות, ומתמלא בסוכר. עייפות אמיתית קיימת; היא פשוט דורשת שעות של מאמץ ולא שש דקות של צנוניות. וכוח רצון כתכונה יציבה לא נבדק כאן כלל.

למה זה ברשימה

כי זו הדוגמה הנקייה ביותר בפקולטה: שלוש שיטות בלתי תלויות הגיעו לאותה מסקנה, ואחת מהן ניבאה אותה מראש.

שנה	2016
תחום	שליטה עצמית
נושא	כוח רצון כמשאב

להעמיק

תשעים ושבעה ניסויים הצליחו במקור, שלושים ושישה שרדו את השחזור

לקריאה

סקירה עברית שמציבה את דלדול האגו בתוך התמונה הרחבה של משבר השחזור.

<https://www.bizportal.co.il/science/news/article/20037206>

המאמר המקורי, 1998

מקורות וארכיון

ניסוי הצנוניות במלואו, כולל ריח העוגיות שנאפו בחדר והפער בין 8.35 ל-20.86 דקות. באנגלית.

[https://faculty.washington.edu/jdb/345/345 Articles/Baumeister et al. \(1998\).pdf](https://faculty.washington.edu/jdb/345/345%20Articles/Baumeister%20et%20al.%20(1998).pdf)

דוח השחזור הרשום, 2016

מקורות וארכיון

עשרים ושלוש מעבדות, 2,141 נבדקים, וגודל אפקט של 0.04 עם רווח סמך שכולל אפס. באנגלית.

https://statmodeling.stat.columbia.edu/wp-content/uploads/2017/11/Hagger_0-407863.pdf

תשובת המחברים המקוריים

מקורות וארכיון

ההתנגדות שלהם למשימה שנבחרה, ולצידה ההודאה שההסכמה לפרוטוקול לא בדוק הייתה טעות. באנגלית.

https://carlsonschool.umn.edu/sites/carlsonschool.umn.edu/files/2019-04/baumeister_vohs_2016_perspectives_comment_on_hagger_rrr_misguided_effort_with_elusive_implications_2_0.pdf

המבחן הרביעית של 2021

מקורות וארכיון

שלושים ושש מעבדות ו-3,531 נבדקים, עם משימות שעונות להתנגדות — ואותה תוצאה. באנגלית.

<https://www.asc.upenn.edu/sites/default/files/2022-06/Vohs et al 2021.pdf>

האמת המרה על סוכר וכוח רצון

מקורות וארכיון

הבדיקה הנפרדת של סיפור הגלוקוז, והמסקנה שאין לו כמעט ערך ראייתי. באנגלית.

https://eprints.lse.ac.uk/109256/1/Bitter_truth_about_glucose_and_willpower_Vadillo_et_al.pdf

ספרים

כוח הרצון

רוי באומייסטר וג'ון טירני · 2011

הספר שמכר את התיאוריה לציבור, כולל הגלוקוז והמילקשייקים, ויצא בעברית בהוצאת מטר. שווה לקרוא כמסמך של הטענה בשיאה, לא כמקור עדכני.

The Seven Deadly Sins of Psychology

2017 · Chris Chambers

מבפנים: איך התמריצים בתחום מייצרים ממצאים כמו 0.62 שמתכווצים ל-0.04, ואיך נולד הפורמט שהכריע כאן. לא אותר תרגום עברי.

Science Fictions

2020 · Stuart Ritchie

משתמש בדלדול האגו כדוגמה מרכזית להטיית פרסום. לא אותר תרגום עברי.

פיאז'ה ומה שנשאר ממנו

1985 · שלושה ילדים, שישה שלבים, ומה קרה כשמדדו תינוקות

ילד בן ארבע רואה שני כוסות מים זהות, ומסכים שיש בהן אותה כמות. שופכים את אחת לכוס צרה וגבוהה, והוא אומר בביטחון מלא שעכשיו יש בה יותר. זו משימת השימור, והיא עובדת. כל הורה יכול לנסות אותה הערב ולקבל את אותה תוצאה. על התופעה הזאת אין ויכוח.

הוויכוח הוא על מה שפיאז'ה בנה עליה. הוא תיאר התפתחות כסדרת שלבים שכל ילד עובר בסדר קבוע, ותיאר את השלב הראשון — השלב החושי-תנועתי, על שישה תת-שלבים — על בסיס שלושה נבדקים: ילדיו שלו. הוא לא דיווח איך נבחרו, כי לא היה מה לדווח.

הבדיקה הראשונה הגיעה מכיוון בלתי צפוי. ב־1974 חזרו על משימת השימור בשינוי אחד: במקום שהמבוגר יסדר מחדש את השורה בכוונה, בובת דובי "שובבה" פיזרה אותה בטעות. יותר ממחצית הילדים בגילאי ארבע עד שש ענו נכון. כלומר הילד הבין את הכמות — הוא פשוט הסיק שאם מבוגר טורח לשנות משהו ואז שואל שוב, כנראה משהו השתנה. זה הבדל בין יכולת לבין ביצוע, והוא מפרק את הלוגיקה של השלבים.

ואז הגיעה שיטת מדידה חדשה: כמה זמן תינוק מסתכל על משהו. ב־1985 נמדדו עשרים ואחד תינוקות בני כחמישה חודשים שראו מסך מסתובב כמו גשר מתרומם, ומאחוריו קופסה. כשהמסך עבר דרך המקום שבו הקופסה אמורה להיות, התינוקות הסתכלו יותר זמן — כאילו ראו משהו בלתי אפשרי. לפי פיאז'ה, ילדים בגיל הזה עדיין אינם יודעים שחפץ ממשיך להתקיים כשהוא נעלם. מכאן צמחה תפיסה שלמה של "ידע ליבה": מערכות מולדות ומצומצמות לחפצים, לסוכנים, למספר ולגיאומטריה — עם מגבלות חדות, כמו יכולת לעקוב אחרי כשלושה חפצים בלבד.

אלא שגם המדידה הזאת נחקרה. כבר ב־1999 נטען שההסתכלות הארוכה משקפת העדפה תפיסתית לתנועה רבה יותר ולא ציפייה מופרת, והוויכוח הזה חי עד היום; מגיניה של השיטה הודו ב־2024 שצריך לשלול את החלופה הזאת מקרה-מקרה. וחשוב מכך: בשנים האחרונות רצו קונסורציומים בינלאומיים ששחזרו ממצאים קלאסיים על תינוקות בעשרות מעבדות בבת אחת. התוצאה מעורבת בצורה מאלפת. בבדיקה הראשונה, על 2,329 תינוקות בשישים ותשע מעבדות, ההעדפה לדיבור מכוון-תינוק אכן אושרה — אבל בגודל של כמחצית ממה שדווח קודם. בבדיקה אחרת, על 521 פעוטות, התוצאה יצאה בכיוון **ההפוך** מהצפוי. ובבדיקה של ממצא מפורסם על תינוקות שמעדיפים דמות שעוזרת על פני דמות שמפריעה — במקור שנים עשר מתוך שנים עשר בחרו בעוזרת — יצא בשלושים ושבע מעבדות ו־567 תינוקות שיעור של 49.34 אחוז. הטלת מטבע.

הסיכום ההוגן קשה יותר משתי הכותרות המתבקשות. פיאז'ה לא "טעה": השימור, שגיאת החיפוש אצל תינוקות, והסדר הכללי של ההתפתחות — כל אלה תופעות אמיתיות וחוזרות. מה שנעלם הוא המבנה: שלבים חדים, אינטליגנציה כדבר אחד שמתפתח כיחידה, וההנחה שאדם מגיע בגיל אחת עשרה לשיא לוגי — מבחנים

שסטודנטים במכללות נכשלים בהם בקביעות. ומה שנשאר, ושכדאי לדעת: הרעיון שילד בונה ידע דרך פעולה ולא מקבל אותו בהסבר, והתובנה ששגיאות של ילדים אינן רעש אלא מידע. גם החוקרים שהפריכו את השלבים עובדים עד היום בתוך המסגרת הזאת.

למה זה ברשימה

כי התופעות שפיאז'ה גילה עדיין קורות בדיוק כפי שתיאר — וההסבר שלו להן כמעט כולו נעלם.

שנה	1985
תחום	התפתחות והתקשרות
נושא	התפתחות קוגניטיבית

להעמיק

ההשפעה המתמשכת של ז'אן פיאז'ה

לקריאה

מה נשאר ומה נעלם, כולל ההודאה שהשלבים החושיים-תנועתיים נשענו על שלושה ילדים. באנגלית.

<https://www.psychologicalscience.org/observer/jean-piaget>

ידע ליבה

לקריאה

סקירה קצרה של ארבע מערכות הידע המולדות, כולל המגבלות החדות שלהן. באנגלית.

<https://www.harvardlds.org/wp-content/uploads/2017/01/SpelkeKinzler07-1.pdf>

מה חמשת אלפים תינוקות מלמדים אותנו

לקריאה

סיכום של קונסורציום השחזורים: מה אושר, מה יצא אפס, ומה יצא הפוך מהצפוי. באנגלית.

<https://www.nature.com/articles/s44271-026-00477-w>

שפת התינוקות: כיצד ילדים לומדים לדבר

לקריאה

מקור עברי על מחקר תינוקות עכשווי ועל מה שאפשר למדוד בגיל שבו אי אפשר לשאול.

<https://davidson.org.il/read-experience/scientificamerican/שפת-התינוקות-כיצד-ילדים-לומדים-לדבר/scientificamerican>

קביעות האובייקט בגיל חמישה חודשים, 1985

מקורות וארכיון

המאמר שהציג את שיטת ההסתכלות ואת מערך הגשר המתרום, עם כל הנתונים. באנגלית.

https://infantcognition.web.illinois.edu/wp-content/uploads/2021/12/baillargeon_spelke_wasserman_1985-Object-permanence-in-5-month-old-infants.pdf

The Psychology of the Child

1966 · Jean Piaget & Bärbel Inhelder

הסיכום הקצר של התיאוריה מפי מחבריה — כדאי לקרוא כדי לראות כמה היא מסודרת וכמה מעט נתונים עמדו מאחוריה. לא יותר תרגום עברי.

What Babies Know

2022 · Elizabeth Spelke

הניסוח המלא של תפיסת ידע הליבה מאת מי שפיתחה אותה. לא יותר תרגום עברי.

The Scientist in the Crib

1999 · Gopnik, Meltzoff & Kuhl

הצגה פופולרית של מחקר התינוקות שהחליף את השלבים, כתובה בידי שלושה מהחוקרים המרכזיים. לא יותר תרגום עברי.

התקשרות

1988 · מעשרים ושש משפחות ועד מבחן אישיות באינטרנט

המחקר שהכול יצא ממנו קטן מכפי שנדמה. מרי איינסוורת' עקבה אחרי עשרים ושש דיאדות של אם ותינוק בבולטימור, בשני גלים בשנות השישים, בביקורי בית של כארבע שעות כל שלושה שבועות — כשבעים ושתיים שעות תצפית לכל משפחה בגל הראשון. הנשים גויסו דרך רופאי ילדים, ואיינסוורת' עצמה ציינה מאוחר יותר שהרופאים "נטו לבחור נשים שעניינו אותם".

מתוך אלפי שעות התצפית הזאת נולדה פרוצדורה בת עשרים דקות שהיא מה שכולם מכירים: "המצב הזר". שמונה שלבים, שתי פרידות ושני איחודים, והסיווג נקבע לפי מה שהילד עושה ****באיחוד**** — לא לפי כמה הוא בכה. זה הבדל שכדאי לדעת, כי הוא מקור לרוב אי ההבנות.

ואז השאלה המעניינת: האם החלוקה הזאת אוניברסלית? ב־1988 אוחדו שלושים ושניים מחקרים משמונה מדינות, 1,990 תינוקות. התוצאה: הסיווג הבטוח הוא השכיח ****בכל**** התרבויות שנבדקו — אבל הפילוגים שונים, ובאופן שמעניין אותנו במיוחד: ההבדלים ****בתוך**** מדינה היו בערך פי אחד וחצי מההבדלים בין מדינות. ישראל תופסת כאן מקום מיוחד. בקיבוצים שבהם נהגה לינה משותפת נמצאו 59 אחוז התקשרות בטוחה, לעומת 75 אחוז בקרב תינוקות ישראלים במעון; ובהשוואה ישירה בתוך הקיבוצים עצמם — 48 אחוז בלינה משותפת מול 80 אחוז בלינה משפחתית. הוויכוח אינו על המספרים אלא על משמעותם: האם זו עדות לכך שטיפול לילי לא עקבי, בידי שומרת מתחלפת, מייצר בדיוק את מה שהתיאוריה חוזה — או שהפרוצדורה עצמה מכוילת תרבותית ומטילה על תינוקות מסוימים מנת לחץ גדולה מדי. שימי לב שמדובר בהשוואה בין 23 ל־25 תינוקות, ועל סידור לילה שהקיבוצים עצמם זנחו.

מכאן לשאלה שמעניינת כל קורא: האם זה מנבא משהו? הכיוון קיים והוא קטן. במטא־אנליזות, הפער בין התקשרות בטוחה ללא בטוחה נותן גודל אפקט של 0.39 לכשירות חברתית, 0.31 להתנהגות מוחצנת, 0.15 למופנמת. זה בסדר הגודל של השפעת הורות על עבריינות — אמיתי, לא מכריע. גם ההעברה בין הדורות נמדדה: בסניתזה של תשעים וחמישה מדגמים ו־4,819 נבדקים, המתאם בין ראיון ההתקשרות של ההורה לסיווג התינוק הוא 0.31, ורגישות הורית מסבירה רק כרבע ממנו. שווה להשוות: במטא־אנליזה קודמת, מ־1995, המספר היה 0.47. האפקט התכווץ ככל שהמדגמים גדלו.

ומה עם יציבות לאורך החיים? כאן הוויכוח חריף. גישה אחת מוצאת מתאם ארוך טווח שמתייצב סביב 0.39; מטא־אנליזה גדולה על 127 מאמרים ו־21,072 מדידות מוצאת 0.39 בממוצע אך בדעיכה ברורה — 0.45 במרווחים של עד חצי שנה, 0.20 במרווחים של חמש עד חמש עשרה שנים, ובמרווחים ארוכים מחמש עשרה שנה 0.14, לא מובהק.

ועכשיו החלק שקשה לשימוע. "סגנון התקשרות" כפי שהוא מוכר היום — חרד, נמנע, בטוח, כטוית אישיות — לא מגיע מאינספור ולא מבולבי. הוא מגיע משאלון שפורסם בעיתון בדנוור ב־1985; החוקרים ניתחו את 620 התשובות הראשונות שהגיעו בדואר. ובדיקה טקסונומטרית על כ־4,700 מבוגרים, בשלוש שיטות שנועדו במפורש לגלות טיפוסים, לא נמצאו טיפוסים: השונות רציפה, על שני צירים, ולא ארבע קופסאות. אפילו אמיר לוין, שכתב את רב המכר שהפיץ את החלוקה, פרסם ב־2026 ספר שמתקן אותה — שהסגנונות ניתנים לשינוי, משתנים בין מערכות יחסים, ואינם נחתמים בגיל שלוש — ומזהיר במפורש מלהתייחס אליהם "כמו מזלות".

למה זה ברשימה

כי בין המחקר המקורי לבין "סגנון ההתקשרות" שאנשים מייחסים לעצמם היום עומדים שני גופי מחקר שונים לגמרי, ושאלון עיתון אחד.

שנה	1988
תחום	התפתחות והתקשרות
נושא	תיאוריית ההתקשרות

להעמיק

המצב הזר וסגנון התקשרות

לקריאה

תיאור מסודר בעברית של הפרוצדורה, של הקטגוריות ושל הפילוג במדגם המקורי.

https://www.betipulnet.co.il/lexicon/התקשרות/המצב_הזר_וסגנון

מה קורה לילדים בלי הורים

לקריאה

רקע עברי על הרלו, בולבי ואינספור ועל השאלה מה קורה כשאין דמות התקשרות יציבה.

<https://davidson.org.il/read-experience/sciencepanorama/מה-קורה-לילדים-בלי-הורים>

מחבר "Attached" מתקן את קוראיו

לקריאה

הסבר עדכני למה סגנונות ההתקשרות אינם קטגוריות קבועות ואינם נחתמים בילדות. באנגלית.

<https://slate.com/life/2026/04/secure-book-attached-anxious-avoidant-attachment-style-theory.html>

טיפול קולקטיבי בקיבוץ והתקשרות

מקורות וארכיון

הנתונים הישראליים המלאים, כולל ההשוואה בין לינה משותפת ללינה משפחתית באותם קיבוצים. באנגלית.

<https://www.marinusvanijzendoorn.nl/wp-content/uploads/2012/07/Aviezer-Van-IJzendoorn-Sagi-et-al-1994.-Psychological-Bulletin.-Collective-early-childcare-kibbutzim.pdf>

לסגור את פער ההעברה

מקורות וארכיון

מטא-אנליזה של שלושה עשורים: ההעברה בין הדורות אמיתית, וקטנה ממה שחשבו. באנגלית.

[https://discovery.ucl.ac.uk/id/eprint/1477395/1/Meta-analysis intergenerational transmission attachment accepted.pdf](https://discovery.ucl.ac.uk/id/eprint/1477395/1/Meta-analysis%20intergenerational%20transmission%20attachment%20accepted.pdf)

ספרים

A Secure Base

1988 · John Bowlby

הניסוח הבהיר ביותר של התיאוריה מפי מי שבנה אותה, ובלי שום קשר לתעשיית השאלונים שצמחה עליה. לא אותר תרגום עברי.

Cornerstones of Attachment Research

2020 · Robbie Duschinsky

ההיסטוריה הארכיונית של התחום: מה בדיוק נמצא בכל מחקר מכונן, ומה נוסף לו בדיעבד. פתוח לקריאה חופשית. לא אותר תרגום עברי.

Attachment in Adulthood

2007 · Mario Mikulincer & Phillip Shaver

הסיכום המחקרי של התקשרות אצל מבוגרים, מאת חוקר ישראלי מרכזי בתחום — והניגוד המדויק לשאלוני הרשת. לא אותר תרגום עברי.

מילגרם

1963 · מה בדיוק נמדד בניסוי הציות

הניסוי בייל, בשנים 1961 ו-1962: נבדק מתבקש להעביר מכות חשמל בעוצמה הולכת וגדלה ל"תלמיד" שטועה בתרגיל זיכרון. המחולל נראה אמיתי, שלושים מתגים, מחמישה עשר וולט ועד ארבע מאות וחמישים, עם תווית "סכנה: מכה חמורה". איש אינו מקבל מכה. במאמר שפורסם ב-1963, עשרים ושישה מתוך ארבעים איש הגיעו עד הסוף. שישים וחמישה אחוז.

המספר הזה הוא אמיתי, והוא גם מטעה. מילגרם הריץ עשרים ושלוש וריאציות על כשבע מאות ושמונים אנשים. כשחוקרים סיכמו לאחרונה עשרים ואחת מהן, על 740 נבדקים, הציות הממוצע היה 43.6 אחוז — והטווח נע מאפס ועד 92.5. הגרסה שבה הנבדק אינו לוחץ בעצמו על המתג הפיקה את השיעור הגבוה; הגרסה שבה הנבדק בוחר בעצמו את העוצמה הפיקה כמעט אפס. כלומר מה שקובע אינו "סמכות" מופשטת אלא כמה בדיוק המערך מכוון את האדם.

ואז מגיע הארכיון. ג'ינה פרי עברה על ההקלטות ועל התיקים ומצאה שהנסיין חרג מהתסריט שוב ושוב, חזר על ההנחיות מחוץ לסדר והמציא חדשות — כלומר ההליך לא היה מתוקנן כפי שהמאמר מציג. וממצא שני, שמשנה את פירוש המספר עצמו: בדיקה שערך עוזרו של מילגרם ומעולם לא פורסמה, שנותרה מחדש ב-2019, מראה שבשמונה עשרה מתוך עשרים ושלוש הווריאציות **מי שהאמין באמת שהוא גורם כאב העביר מכות חלשות יותר** ממי שלא האמין. תפיסת הכאב הגבירה מרי; הספקנות הגבירה ציות. חלק ממה שנספר כ"ציות" הוא חשד שזה משחק.

גם הצד האתי, שבגללו הניסוי נחשב לאבן דרך בהגנה על נבדקים, שונה מהסיפור. לפי הארכיון, שלושה רבעים מהנבדקים יצאו מהמעבדה בלי שתוחקרו, וחלקם קיבלו את האמת בדואר חודשים ואף שנה אחר כך.

ולמרות הכול — התופעה שרדה. ב-2009 נערך שחזור חלקי: עוצרים ב-150 וולט, כי אצל מילגרם 79 אחוז ממי שעבר את הנקודה הזאת המשיך עד הסוף. שבעים אחוז מתוך שבעים נבדקים המשיכו מעבר לה, לעומת 82.5 אחוז אצל מילגרם — הפרש לא מובהק, ארבעים וחמש שנה אחר כך, גם אצל נשים וגם אצל גברים. יש גם ביקורת על השחזור הזה, שסינון הנבדקים מנפח את התוצאה. ובכל זאת, אין כאן נתונים מומצאים ואין הפרכה.

מה שכן השתנה הוא ההסבר. ארבע ההנחיות שהנסיין היה אמור לומר מסודרות בסולם עולה של לחץ, והרביעית היא היחידה שהיא פקודה של ממש: "אין לך ברירה, אתה חייב להמשיך". מתברר שהיא הייתה **הכי פחות** יעילה, ולרוב הובילה לסירוב; היעילה ביותר הייתה השנייה, "הניסוי מחייב שתמשיך". מכאן הפרשנות שרווחת היום: אנשים לא צייתו בעיוורון אלא שיתפו פעולה עם מפעל שנראה להם ראוי, ופרשו ברגע שהופעל עליהם כוח.

ולבסוף, ההקשר שמילגרם עצמו פתח בו את המאמר — השואה. היסטוריונים אינם מקבלים את ההשוואה. ביוזפוב, ביולי 1942, הציע מפקד גדוד משטרה גרמני לכחמש מאות אנשיו להשתחרר מירי בכאלף וחמש מאות יהודים; כשנים עשר נענו מיד, ובאומדנים מאוחרים כעשרה עד עשרים אחוז ביקשו העברה בסופו של דבר. כלומר יותר משמונים אחוז המשיכו לירות — בתנאים של קרבה לקורבן, מגע פיזי, פיקוח רופף ושיקול דעת, שהם בדיוק התנאים שבווריאציות של מילגרם **הגבירו** סירוב. המודל שלו מנבא כאן את ההפך ממה שקרה.

למה זה ברשימה

כי המספר שכולם מכירים מגיע מתנאי אחד מתוך עשרים ושלושה, והפקודה היחידה בכל התסריט הייתה גם הכי פחות יעילה.

שנה	1963
תחום	פסיכולוגיה חברתית
נושא	ציות לסמכות

להעמיק

מטא־מילגרם

לקריאה

ניתוח של עשרים ואחת מהווריאציות: ממוצע 43.6 אחוז, בטווח שנע מאפס ועד 92.5. באנגלית.

<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0093927>

אמינות וספקנות בניסויי מילגרם

לקריאה

הניתוח מחדש של בדיקה שלא פורסמה: מי שהאמין שהוא גורם כאב ציית פחות. באנגלית.

<https://journals.sagepub.com/doi/10.1177/0190272519861952>

מה הניסויים של מילגרם וזימברדו באמת מראים

לקריאה

הפרשנות שהחליפה את "ציות עיוור", כולל הממצא על ההנחיה הרביעית. באנגלית.

<https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1001426>

לשפוך את מילגרם עם המים

לקריאה

מסה בעברית שמודה בכל הפגמים ובכל זאת מסרבת לזרוק את הממצא — בדיוק העמדה שהיחידה מגיעה אליה.

<https://alaxon.co.il/article/מילגרם-עם-המים-לשפוך-את-מילגרם-עם-המים>

המאמר המקורי, 1963

מקורות וארכיון

שמונה עמודים שבהם אפשר לראות מה דווח, מה לא, ואיך נולד המספר שישים וחמישה אחוז. באנגלית.

<https://www.columbia.edu/cu/psychology/terrace/w1001/readings/milgram.pdf>

האליבי של הציות

מקורות וארכיון

הבחינה ההיסטורית של השוואת מילגרם לשואה, עם המקרה של יוזפוב. באנגלית.

<https://psychologyrocks.org/wp-content/uploads/2016/11/obedience-alibi-article-2-1.pdf>

ספרים

Obedience to Authority

1974 · Stanley Milgram

הספר שבו מילגרם מציג את כל הווריאציות ולא רק את המפורסמת, ומנסח את התיאוריה שלו. לא אותר תרגום עברי.

Behind the Shock Machine

2012 · Gina Perry

העבודה הארכיונית שעליה נשען החלק השני של היחידה: מה ההקלטות והתיקים מראים על מה שקרה בפועל. לא אותר תרגום עברי.

Ordinary Men

1992 · Christopher Browning

המקור ההיסטורי ליוזפוב, ולשאלה למה מסגרת הציות אינה מסבירה את מה שקרה שם. לא אותר תרגום עברי.

כלא סטנפורד

1971 · מה ההקלטות מראים על מה שקרה במרתף

באוגוסט 1971 הוסב מרתף המחלקה לפסיכולוגיה בסטנפורד לכלא מאולתר. סטודנטים חולקו באקראי לסוהרים ולאסירים. התכנון היה לשבועיים; הניסוי הופסק אחרי שישה ימים. הגרסה שנכנסה לכל ספר לימוד היא זו שכתוב באתר הרשמי של הניסוי עד היום: "תוך ימים ספורים הסוהרים שלנו הפכו סדיסטים והאסירים דיכאוניים". המסקנה: המצב מנצח את האישיות, ואדם רגיל שמקבל תפקיד אכזרי יהפוך אכזרי.

ב־2018 וב־2019 נפתחו הארכיונים וההקלטות, ומה שיש בהם משנה את המשפט הזה מהיסוד.

ראשית, הסוהרים לא "הפכו" לכלום — הם תודרכו. בהקלטה מ־14 באוגוסט 1971, יום לפני תחילת הניסוי, אומר זימברדו לסוהרים שהם יוכלו ליצור שעמום, ליצור פחד, ליצור תחושה של שרירותיות, ומוסיף: "יש לנו כוח מוחלט במצב הזה. להם אין שום כוח". הסוהר הראשי נתן הנחיות טקטיות למסדרים, ובהן שתיקה כפויה, והסביר שסרקזם הוא כלי השפלה "יעיל מאוד". שנית, מבנה הניסוי כולו לא היה מקורי: אחד עשר מתוך שבעה עשר הכללים הועתקו מסימולציית כלא סטודנטיאלית שאותו סוהר ראשי הריץ שלושה חודשים קודם לכן — עובדה שלא פורסמה במשך כארבעים שנה. ושלישית, הסוהרים כלל לא ידעו שהם הנחקרים; הם הובלו להאמין שהם עוזרי מחקר.

ואז ההקלטות עצמן. הסוהר שהתפרסם כאכזרי ביותר אמר בתחקיר, ב־20 באוגוסט 1971: "לאורך כל הניסוי הייתי שחקן, והגזמתי בכוונה... שיחקתי תפקיד". סוהר אחר כתב במכתב: "תמיד שיחקתי". וסוהר שלישי דיווח שהסוהר הראשי או זימברדו "הנחו אותי במפורש להתנהג בצורה מסוימת". האסיר שהתמוטטותו היא הרגע הרגשי של כל הסיפור אמר בראיון, יום לפני סיום הניסוי: "החלטתי שאני רוצה לצאת... הבנתי שאתה מעביר אותי סיפורים... הדרך הקלה ביותר הייתה פשוט לשחק משוגע... בניתי את זה בכוונה".

ולבסוף, הנתונים. מתוך כמאה וחמישים שעות ניסוי תועדו כשש שעות וידאו וחמש עשרה שעות אודיו — פחות מחמישה עשר אחוז — וביום השלישי לא נאסף דבר. אחד מעוזרי המחקר כתב בדוח פנימי שהמדגם המצולם היה "מוטה ברצינות", כי צילמו בעיקר את הדרמטי והחריג. לא היו השערות שנוסחו מראש, לא הייתה קבוצת ביקורת, לא היו מדדים מתוקפים ולא נערכו מבחנים סטטיסטיים. הודעה לעיתונות שיצאה ביום השני הצהירה שהמטרה היא להביא למודעות את הצורך ברפורמה בבתי הכלא. המסקנה הוגדרה לפני שהיו נתונים.

זימברדו השיב לביקורת. הוא מודה בפגם המרכזי — שלא היה משקיף מדעי עצמאי ושהוא עצמו הפך למנהל הכלא — אך טוען שבסרטי היום הראשון הסוהרים דווקא מצחקקים, ושההקשחה באה רק אחרי מרד ביום השני; ולגבי האסיר שהתמוטט הוא כותב שחוקר שמאמין שנבדק מתמוטט מחויב אתית להתייחס לכך כאמיתי, גם אם יתברר אחר כך שהיה מדומה.

מה שכן אפשר ללמוד? יש ניסוי אחר. ב־2001 ערכו שני חוקרים בריטים סימולציה דומה, בליווי תיעוד מלא ופיקוח אתי: חמישה עשר גברים, שמונה ימים. הסוהרים **לא** התלכדו ונרתעו מלהפעיל סמכות; האסירים פיתחו זהות קבוצתית והפילו את משטר הסוהרים ביום השישי. ואז הוקמה קומונה שוויונית — והיא התפרקה, ובעקבות הכישלון שלה הציעו ארבעה משתתפים להקים משטר נוקשה בהרבה, ובשלב הזה הופסק המחקר. המסקנה שלהם: עריצות מתחילה להיראות אטרקטיבית כשקבוצה נכשלת בלנהל את עצמה — לא כשנותנים לאנשים מדים.

חשוב לא לגלוש לקצה השני. ההשפלות, שלילת השינה והשקיות על הראש קרו באמת; שניים מהאסירים התמוטטו באמת; וזימברדו עצמו איבד פרספקטיבה עד שקולגה עימתה אותו. מה שהביקורת מפילה היא הסיביות, לא האירועים. התפקיד לא פעל "לבדו" — מישו ספק אותו, הצדיק אותו והעניק לו רשות.

למה זה ברשימה

כי הניסוי המצוטט ביותר על טבע האדם לא היה ניסוי, ומה שהוא כן הראה מעניין יותר ממה שסיפרו עליו.

שנה	1971
תחום	פסיכולוגיה חברתית
נושא	תפקיד ואכזריות

להעמיק

תגובתו של זימברדו לביקורת

לקריאה

הודאה בפגם אחד מרכזי, והתעקשות על כל השאר — כולל טיעון אתי מעניין לגבי התמוטטות מדומה. באנגלית.

<https://www.prisonexp.org/response>

ניסוי הכלא המפורסם: לא מה שחשבתם

לקריאה

הסיכום העברי הטוב ביותר של הביקורת, כולל ציטוט מהתדריך המוקלט.

<https://davidson.org.il/read-experience/reasonabledoubt/ניסוי-הכלא-המפורסם-לא-מה-שחשבתם>

תוחלת החיים של שקר

לקריאה

הכתבה מ־2018 שהביאה לראשונה את האסיר ואת הסוהר לדבר, ופתחה את כל הפרשה. באנגלית.

<https://gen.medium.com/the-lifespan-of-a-lie-d869212b1f62>

לפרק את ניסוי הכלא של סטנפורד

מקורות וארכיון

המאמר שקבע את הדיון: מה ההקלטות והמסמכים בארכיון מראים על התהליך, על ההעתקה ועל היקף התייעוד. באנגלית.

<https://gwnet.net/doc/psychology/2019-letexier.pdf>

האתר הרשמי של הניסוי

מקורות וארכיון

הגרסה שזימברדו עצמו מספר, יום אחר יום. כדאי לקרוא אותה לפני הביקורת, כדי לראות מה בדיוק נטען. באנגלית.

<https://www.prisonexp.org/>

מחקר הכלא של הבייביסי

מקורות וארכיון

הסימולציה המתועדת שיצאה הפוך: הסוהרים לא התלכדו, האסירים ניצחו, ואז דווקא הכישלון הוליד קריאה למשטר נוקשה. באנגלית.

[http://www.bbcprisonstudy.org/pdfs/bjsp\(2006\)tyranny.pdf](http://www.bbcprisonstudy.org/pdfs/bjsp(2006)tyranny.pdf)

ספרים

The Lucifer Effect

2007 · Philip Zimbardo

הניסוח המלא של הטיעון מפי זימברדו, כולל ההשוואה לאבו־גרייב — וכדאי לקרוא אותו מול המאמר מ־2019. לא אותר תרגום עברי.

Histoire d'un mensonge

2018 · Thibault Le Texier

החקירה הארכיונית המלאה, בצרפתית, שממנה נגזר המאמר האנגלי. לא אותר תרגום עברי.

Ordinary Men

1992 · Christopher Browning

מועיל גם כאן: מה באמת קרה כשאנשים רגילים קיבלו אפשרות לסרב. לא אותר תרגום עברי.

אפקט המתבונן

1964 · הכתבה שיצרה תופעה, והתופעה שנמצאה בכל זאת

ב 27 במרץ 1964, שבועיים אחרי רצח קיטי ג'נובזה בקווינס, פרסם ה"ניו יורק טיימס" כתבה תחת הכותרת "37 שראו רצח לא התקשרו למשטרה". בגוף הכתבה נכתב 38. הטענה: שלוש התקפות על פני חצי שעה, עשרות אנשים שצפו מהחלונות, ואיש לא עשה דבר. הכתבה נולדה בשיחת צהריים: מפקד משטרת ניו יורק אמר לעורך שהסיפור מקווינס "הוא אחד לספרים", והעורך שלח כתב.

כמעט כל פרט בסיפור הזה אינו מדויק. בעדויות במשפט התבררו שתי התקפות ולא שלוש, והשנייה התרחשה בחדר מדרגות פנימי שכמעט אי אפשר היה לראות לתוכו. שכן פתח חלון בזמן ההתקפה הראשונה וצעק "עזוב את הבחורה!" — והתוקף הסתלק וחזר רק כעשר דקות אחר כך. לפחות שני אנשים התקשרו למשטרה. ושכנה בת שבעים ירדה לחדר המדרגות, מצאה את ג'נובזה ואחזה בה עד שהאמבולנס הגיע; היא הייתה בחיים כשהמשטרה הגיעה. ב־2016, שנים עשר וחמישים שנה אחרי, כתב העיתון עצמו — בהספד לרוצח — שהכתבה מ־1964 "הפריזה באופן גס במספר העדים ובמה שהם תפסו".

שימי לב למה שאי אפשר לעשות כאן: להחליף את המספר השגוי במספר נכון. אין כזה. סגן התובע אמר שמצאו "בערך חצי תריסר" עדים שימושיים; מקורות אחרים מדברים על כתריסר ששמעו או ראו משהו. כל מספר חלופי שמשהו יציג הוא מיתוס חדש.

ומכאן החלק המעניין: הכתבה השגויה הולידה מחקר אמיתי. ב־1968 בנו דארלי ולטנה ניסוי שבו נבדק שומע התקף אפילפסיה דרך אינטרקום. מי שחשב שהוא היחיד שמאזין דיווח ב־85 אחוז מהמקרים; מי שחשב שיש עוד מאזין אחד — ב־62; מי שחשב שיש ארבעה — ב־31. גם זמני התגובה התארכו בהתאם. בניסוי אחר, חדר שהתמלא בעשן: שבעים וחמישה אחוז דיווחו כשהיו לבד, ורק עשרה אחוזים כשישבו לידם שני שותפים אדישים.

האפקט הזה עמד במבחן. במטא־אנליזה מ־2011, על 105 גדלי אפקט ויותר משבעת אלפים נבדקים, הוא קיים ובגודל קטן עד בינוני. ובאותה מטא־אנליזה נמצא גם התנאי שמחליש אותו ואף הופך אותו: כשהמצב נתפס כמסוכן, כשהתוקף נוכח, וכשהמחיר של ההתערבות הוא פיזי — האפקט נחלש, והחוקרים מסבירים שמצב מסוכן מזוהה מהר וברור יותר כחירום ולכן מייצר עוררות והתערבות.

וב־2019 נעשתה הבדיקה שאי אפשר לעשות במעבדה. חוקרים אספו 219 קטעי מצלמות אבטחה של עימותים אלימים אמיתיים ברחוב — באמסטרדם, בקייפטאון ובלנקסטר — וספרו. לפחות אדם אחד התערב ב־90.9 אחוז מהמקרים, בממוצע 3.76 מתערבים לעימות. ולא רק זה: **כל** צופה נוסף העלה במעט את הסיכוי שמישהו יתערב. גם בקייפטאון, האתר עם שיעור הפשיעה הגבוה ביותר, לא נמצא הבדל.

איך שתי התוצאות מתיישבות? כנראה כך: הסיכוי שאדם **מסוים** יתערב אכן יורד ככל שיש יותר אנשים, והסיכוי ש**מישהו** יתערב עולה. שני המשפטים נכונים בו זמנית, והבלבול ביניהם הוא כל העניין. מה שנפל אינו אפקט המתבונן אלא "אדישות עירונית" – התמונה של עיר שבה איש אינו נוקף אצבע. הרחוב, כשמדדו אותו, מתנהג אחרת.

למה זה ברשימה

כי הסיפור שממנו נולד המחקר היה שגוי, הממצא שנבנה עליו נכון, והמציאות ברחוב הפוכה משניהם.

שנה	1964
תחום	פסיכולוגיה חברתית
נושא	התפזרות אחריות

להעמיק

תורת הקבוצות

לקריאה

פתיחה עברית קצרה על אפקט הצופה מהצד ועל התפזרות האחריות, שגם מציינת שהדיווח העיתונאי על הרצח הוטל בספק.

<https://davidson.org.il/read-experience/askexpert/social-group-theory/>

בצד הבטוח של הדברים

לקריאה

מסה בעברית על הדרך שבה "העומד מן הצד" הפך מדמות תמימה לשותף לאשמה, עם רצח ג'נובה במרכז.

<https://alaxon.co.il/article/הדברים-של-הבטוח-בצד/>

ההספד שבו העיתון הודה

לקריאה

הטקסט שבו ה"ניו יורק טיימס" כותב בעצמו שהכתבה מ־1964 הפריזה באופן גס. באנגלית.

<https://www.bostonglobe.com/metro/obituaries/2016/04/09/winston-moseley-killer-kitty-genovese-dies-prison/FxsOVdAgiTaQhF2n702WCN/story.html>

האמת מאחורי סיפור קיטי ג'נובה

לקריאה

סיכום קצר של מה שנמצא בבדיקה ההיסטורית, כולל דבריו של סגן התובע. באנגלית.

<https://www.bps.org.uk/research-digest/truth-behind-story-kitty-genovese-and-bystander-effect>

מה קורה באמת ברחוב

מקורות וארכיון

המחקר על 219 עימותים אמיתיים במצלמות אבטחה בשלוש ערים, והתערבות ב־90.9 אחוז מהם. באנגלית.

<https://gwnet.net/doc/crime/2019-philpot.pdf>

ספרים

Kitty Genovese: The Murder, the Bystanders, the Crime That Changed America

2014 · Kevin Cook

השחזור העיתונאי המסודר של מה שקרה באותו לילה ושל מה שהעיתון עשה ממנו. לא אותר תרגום עברי.

No One Helped

2015 · Marcia M. Gallo

טוען שהעיתון יצר את המיתוס, ומשחזר את מה שנחקק ממנו. לא אותר תרגום עברי.

The Unresponsive Bystander

1970 · Bibb Latané & John Darley

הספר של שני החוקרים שהפכו את הכתבה למחקר של ממש. לא אותר תרגום עברי.

תנוחות כוח

2010 · ארבעים ושניים נבדקים, שבעים ושבעה מיליון צפיות

המחקר פורסם ב־2010. ארבעים ושניים נבדקים החזיקו שתי תנוחות, דקה כל אחת. שתי דקות בסך הכול. הטענה בתקציר: מי שהחזיק תנוחה פתוחה וגדולה חווה עלייה בטסטוסטרון, ירידה בקורטיזול, תחושת כוח מוגברת ונכונות רבה יותר לקחת סיכון; מי שהחזיק תנוחה מכווצת — ההפך. מדד נטילת הסיכון היה זה: נותנים לנבדק שני דולר, והוא יכול לשמור אותם או להטיל קובייה ולזכות בארבעה בסיכוי חצי.

ב־2012 עלתה הרצאת טד שהפכה את זה לתופעה עולמית. נכון להיום היא צברה יותר משבעים ושבעה מיליון צפיות — מהנצפות בכל הזמנים.

ב־2015 נעשה שחזור על 200 נבדקים, עם שיפורים: התנוחות הוחזקו שלוש דקות במקום אחת, וההוראות ניתנו במחשב כך שהבודק היה עיוור לתנאי. התוצאה חדה: תחושת הכוח הסובייקטיבית **כך** שוחזרה, בגודל אפקט 0.34. הטסטוסטרון, הקורטיזול ונטילת הסיכון — לא. ולא רק שלא היו מובהקים; המגמות הלא־מובהקות הלכו בכיוון ההפוך לחיזוי.

ואז קרה דבר נדיר. בספטמבר 2016 פרסמה דנה קארני, החוקרת הראשית של המחקר המקורי, מסמך קצר בעמוד הסגל שלה בברקלי. היא כותבת: "אינני מאמינה שאפקטים של תנוחות כוח אמיתיים"; "אינני חוקרת את האפקטים הגופניים של תנוחות כוח"; "אני מרתיעה אחרים מלחקור אותן". והיא ממשיכה ומפרטת מה נעשה בנתונים: "גודל המדגם זעיר"; "הנתונים רופפים. האפקטים קטנים ובקושי קיימים"; המשתנה המדווח־עצמי "עבר סינון, במובן זה שנשאלו הרבה שאלות על כוח ונבחרו אלה שעבדו"; ולגבי נטילת הסיכון — "ערך אחד היה 0.052 ואחר 0.05, ודווח הקטן מביניהם". חמישה נבדקים הוצאו מהניתוח.

שימי לב שקארני אינה מתארת הונאה, והיא אומרת זאת במפורש: אלה היו נהלים מקובלים ולגיטימיים בעיני כולם ב־2010. זה בדיוק מה שהופך את המסמך הזה לשימושי כל כך — הוא מראה איך נראה מחקר שגוי שנעשה ביושר.

הוויכוח לא נגמר שם. ב־2017 נבדקה כל ספרות תנוחות הכוח בכלי סטטיסטי שמזהה סינון תוצאות, והמסקנה הייתה שהעוצמה הסטטיסטית הממוצעת של כל הספרות הזאת היא בסביבות חמישה אחוז. שנה אחר כך פרסמה איימי קאדי, שותפה למחקר המקורי ומי שנשאה את ההרצאה, ניתוח נגדי על חמישים וחמישה מחקרים — ומצאה ערך ראיתי ברור. אבל שווה לקרוא מה בדיוק היא מצאה ערך ראיתי **עבורו**: רגש, תחושה והערכה עצמית. עבור מדדים התנהגותיים והורמונליים היא מדווחת בעצמה על "היעדר ערך ראיתי". כלומר שני הצדדים מסכימים היכן זה נגמר. מטא־אנליזה מאוחרת יותר, על שבעים ושלושה מחקרים ויותר משבעת אלפים נבדקים, הציעה שמה שקיים נובע בעיקר מהימנעות מתנוחה מכווצת ולא מאימוץ תנוחה פתוחה.

ושתי הערות של דיוק. המאמר מ־2010 מעולם לא נמשך והוא עומד בכתב העת עד היום; ההודאה הייתה אישית, בעמוד אינטרנט. ואת ההודאה פרסמה קארני, לא קאדי — קאדי ממשיכה להגן על טענת המשוב הגופני, ובלבול בין השתיים הוא הטעות הנפוצה ביותר בכתובה על הפרשה, בעברית ובאנגלית כאחד.

למה זה ברשימה

כי כאן החוקרת הראשית עצמה פרסמה בדיוק מה עשתה בנתונים — וזה המסמך הכן ביותר שיש בפקולטה הזאת.

שנה	2010
תחום	פסיכולוגיה חברתית
נושא	משוב גופני

להעמיק

ההרצאה שהפיצה את הרעיון

לצפייה

כדאי לצפות ולשים לב לכותרת: היא כוללת היום את המילה "עשויה". באנגלית, עם כתוביות.

https://www.ted.com/talks/amy_cuddy_your_body_language_may_shape_who_you_are

עמדתי בנוגע לתנוחות כוח

לקריאה

שני עמודים שבהם החוקרת הראשית מפרטת בדיוק מה נעשה בנתונים ולמה אינה מאמינה עוד באפקט. באנגלית.

[https://faculty.haas.berkeley.edu/dana_carney/pdf_My position on power poses.pdf](https://faculty.haas.berkeley.edu/dana_carney/pdf_My%20position%20on%20power%20poses.pdf)

מאמר השחזור, 2015

לקריאה

מאתיים נבדקים: תחושת הכוח שוחזרה, ההורמונים ונטילת הסיכון לא. באנגלית.

https://faculty.haas.berkeley.edu/dana_carney/ranehill.pdf

עשור של תנוחות כוח: איפה אנחנו עומדים

לקריאה

סקירה מסודרת של כל הפרשה, מהמחקר המקורי ועד מטא־אנליזה של שבעים ושלושה מחקרים. באנגלית.

<https://www.bps.org.uk/psychologist/decade-power-posing-where-do-we-stand>

עמדת כוח, עמדת חולשה

לקריאה

הדיווח העברי מ־2011 על המחקר המקורי, לפני שנבדק — מסמך משמר של מה שנטען ובאיזה ביטחון.

<https://www.hayadan.org.il/reverse-body-language-0507113>

מה נשאר מהכלכלה ההתנהגותית

לקריאה

כתבה עברית שמציבה את הפרשה בהקשר הרחב ומייחסת נכונה את ההודאה לקארני ולא לקאדי.

<https://www.bizportal.co.il/science/news/article/20037206>

ספרים

Presence

2015 · Amy Cuddy

הספר שיצא מההרצאה, תורגם לשלושים ושתיים שפות ויצא בעברית בשם "נוכחות". מסמך של הטענה בשיאה.

Science Fictions

2020 · Stuart Ritchie

ההסבר הנגיש ביותר למה זה "סינון תוצאות" ואיך הוא מייצר ספרות שלמה מכלום. לא אותר תרגום עברי.

The Seven Deadly Sins of Psychology

2017 · Chris Chambers

על התמריצים שהפכו נהלים כמו אלה שקארני מתארת לסטנדרט מקובל. לא אותר תרגום עברי.

חמשת הגורמים ומיירס-בריגס

1990 · שני מבחני אישיות, ורק אחד מהם נבדק

מודל חמשת הגורמים לא הומצא בישיבה תיאורטית — הוא נשלף ממילון. ב־1936 עברו שני חוקרים על מילון ובר וחילצו כשמונה עשר אלף מילים שבהן אנשים מתארים זה את זה; כארבעת אלפים וחצי מהן סווגו כתכונות יציבות. כשצמצמו את הרשימה והריצו עליה ניתוח גורמים, שוב ושוב התקבלו חמישה: מוחצנות, נעימות, מצפוניות, נירוטיות ופתיחות. ההיסטוריה של המודל היא רשימת חוקרים שקיבלו את אותן חמש קבוצות בנפרד, בשנות הארבעים, החמישים והשישים. במחקר מ־1990 נבדקה היציבות הזאת בארבעה מדגמים בלתי תלויים, ומקדמי ההתאמה ביניהם נעו בין 0.86 ל־0.94 — ואף גורם נוסף לא חזר על עצמו.

גם היציבות לאורך זמן נמדדה, במטא־אנליזה של 152 מחקרי אורך על כ־55 אלף נבדקים: דירוג יחסי בין אנשים מתייצב סביב 0.5 בגיל הקולג', כ־0.6 בגיל שלושים, וכ־0.75 בין גיל חמישים לשבעים.

מול זה עומד מבחן מיירס-בריגס. אותו בנו קתרין בריגס ובתה איזבל מיירס, אף אחת מהן לא פסיכולוגית בהכשרתה, על בסיס טיפוס היונג. איזבל מיירס למדה מדעי המדינה ורכשה את הכלים הפסיכומטריים בהתלמדות אצל איש כוח אדם. וקריטריון האמת שלהן היה אישי במובן המילולי: הן זיהו אנשים שלדעתן מייצגים כל טיפוס, "מתוך היכרות ארוכה ומתוך עשרים שנות תצפית קפדנית", וכתבו פריטים שמפרידים בין אותם אנשים.

שלושה ממצאים מספיקים. ראשית, יציבות: לפי הספרות, בין שלושים ותשעה לשבעים ושישה אחוז מהנבדקים מקבלים טיפוס אחר במבחן חוזר אחרי חמישה שבועות בלבד; על פני תשעה חודשים, כשמונים ושלושה אחוז מהדיכטומיות הבודדות נשמרות, אבל רק כמחצית מהאנשים שומרים על אותו צירוף של ארבע אותיות. השוויון את זה ל־0.6 של חמשת הגורמים על פני כמעט שבע שנים. שנית, טיפוסים: הציונים מתפלגים בצורת פעמון עם שיא במרכז, ולא בשני שיאים כפי שתורת טיפוסים מחייבת. כשהופעלו על הנתונים שיטות סטטיסטיות שנועדו במפורש לאתר טיפוסים חבויים — לא נמצאה שום תמיכה בקיומם. ושלישית, ניבוי: אין ראיות לקשר בין הטיפוס להצלחה בעיסוק.

ובכל זאת: כשני מיליון העברות בשנה, שמונים ותשעה מתוך מאה התאגידים הגדולים, כעשרים מיליון דולר הכנסה שנתית.

שתי הערות שמחייבות הגינות. הראשונה: המבחן אינו מודד כלום — לא נכון. הפריטים שלו אכן מתואמים עם כמה מממדי חמשת הגורמים. מה שנכשל הוא הקופסאות, לא השאלות. וזה גם לקח סטטיסטי כללי: כשחותכים משתנה רציף לשניים מאבדים כוח סטטיסטי שקול לזריקה של כשלושים ושמונה אחוז מהנבדקים. השנייה: אם מבחן כזה מרגיש למישהו מדויק להפליא, ההסבר אינו טיפשות אלא נטייה מוכרת לקבל תיאורים כלליים וחיוביים כאילו נכתבו עלינו אישית.

ולבסוף, כדי לא ליפול לצד השני: גם חמשת הגורמים אינם קסם. במבנה הסקרים הקיים הוא אינו משחזר את עצמו היטב מחוץ לעולם המערבי — במחקר על 94,715 נשאלים בעשרים ושלוש אוכלוסיות שאינן מערביות נמצא שהמבנה אינו מחזיק ברוב סקרי השטח, ושבאזורים מסוימים זהות המראיין הסבירה עד רבע מהשונות. וכוח הניבוי שלו צנוע: מצפוניות מנבאת ביצועים בעבודה במתאם של כ-0.19, ואריכות ימים במתאם של כ-0.11. אמיתי, נמדד, קטן.

למה זה ברשימה

כי שמונים ותשעה מתוך מאה התאגידים הגדולים בעולם משתמשים במבחן שבו עד שלושה מכל ארבעה אנשים מקבלים טיפוס אחר אחרי חמישה שבועות.

שנה	1990
תחום	אישיות ואינטליגנציה
נושא	מדידת אישיות

להעמיק

האם תיאוריית מיירס-בריינס תקפה

לקריאה

סקירה שבודקת אחת-אחת את הנחות היסוד של המבחן, כולל היעדר שני שיאים בהתפלגות והיקפו המסחרי. באנגלית.

https://swanpsych.com/publications/SteinSwanMBTITheory_2019.pdf

על גרפולוגיה והונאה עצמית

לקריאה

הסבר עברי למנגנון שגורם לנו להרגיש שאבחון אישיות חסר תוקף "מדויק לגמרי".

<https://davidson.org.il/read-experience/reasonabledoubt/על-גרפולוגיה-והונאה-עצמית>

תיאור חלופי של האישיות: מבנה חמשת הגורמים

מקורות וארכיון

המאמר שבו התקבע המודל, עם מקדמי ההתאמה בין ארבעה מדגמים בלתי תלויים. באנגלית.

https://projects.ori.org/lrg/pdfs_papers/goldberg.big-five-factorsstructure.jpssp.1990.pdf

תיקון אומדני התוקף בבחירת עובדים

מקורות וארכיון

המאמר שהוריד את המצפוניות מ-0.31 ל-0.19 ואת היכולת הקוגניטיבית מ-0.51 ל-0.31. באנגלית.

<https://gwern.net/doc/statistics/meta-analysis/2021-sackett.pdf>

האם אנשים מצפוניים חיים יותר

מקורות וארכיון

מטא-אנליזה של עשרים מדגמים: קשר אמיתי, וצנוע, של 0.11. באנגלית.

http://www.peggykern.org/uploads/5/6/6/7/56678211/kern_2008_conscientiousness.pdf

ספרים

The Personality Brokers

2018 · Merve Emre

ההיסטוריה הארכיונית של בריגס ומיירס ושל הדרך שבה מבחן תוצרת בית הפך לתעשייה. לא אותר תרגום עברי.

Personality: What Makes You the Way You Are

2007 · Daniel Nettle

מבוא קצר וקריא לחמשת הגורמים, לקורא שמתחיל מאפס. לא אותר תרגום עברי.

Me, Myself, and Us

2014 · Brian R. Little

על כך שתכונות הן רציפות ותלויות מצב, ולא קופסאות. לא אותר תרגום עברי.

מנת משכל

2018 · מה המבחן מנבא, ומה אי אפשר להסיק ממנו

נתחיל במה שנמדד. מבחני מנת משכל מנבאים ציונים בבית הספר במתאם של כ-0.5, ומספר שנות לימוד במתאם של כ-0.55. ביצועים בעבודה — כאן צריך לדייק, כי המספר השתנה: במשך שנים צוטט 0.51, ובעבודה מ-2022 שהראתה שתיוקוני הטווח הקודמים היו מוגזמים ירד האומדן ל-0.31. הכנסה: 0.23 בלבד, במטא-אנליזה של מחקרי אורך על כ-29 אלף איש — מספר שהחוקר עצמו כינה "נמוך עד מאכזב", אף שאף מנבא אחר לא עושה שם טוב יותר.

הממצא החזק ביותר בכל היחידה הוא דווקא אחר. ב-1 ביוני 1932 נבחנו כמעט כל ילדי סקוטלנד בני אחת עשרה ביום אחד. שישים וחמש שנים אחר כך אותרו 2,230 מילדי אברדין ונבדק מי עדיין בחיים בגיל שבעים ושש. על כל חמש עשרה נקודות נמוכות יותר בגיל אחת עשרה, הסיכון היחסי להישרדות היה 0.79. מבחן בכיתה, בילדות, שמנבא תמותה שישה עשרים אחר כך.

ועכשיו המונח שגורר את רוב הבלבול: תורשתיות. אומדני התורשתיות של מנת משכל נעים בין ארבעים לשמונים אחוז, והם עולים עם הגיל. אלא שתורשתיות אינה "כמה מהאינטליגנציה שלך באה מגנים". זהו מדד על **שונות באוכלוסייה מסוימת, בטווח הסביבות שקיים בה**.* דוח הוועדה של האיגוד הפסיכולוגי האמריקני מנסח זאת חד: "טעות נפוצה היא להניח שמשום שמשוהו תורשתי הוא בהכרח בלתי ניתן לשינוי. זה שגוי". הדוגמה שלהם מצוינת — אוצר מילים הוא תורשתי מאוד, וכל מילה בו נלמדה ממישהו. ומחקר ידוע הראה שבמשפחות במצב חברתי-כלכלי נמוך כמעט כל השונות מוסברת בסביבה וכמעט לא בגנים, ובמשפחות מבוססות ההפך — כלומר גם המספר עצמו משתנה לפי התנאים.

וכאן נכנסת העובדה שמכריעה את הדיון הקשה ביותר בתחום. במשך המאה העשרים עלו ציוני מנת המשכל בכל מדינה שנמדדה, בכשלוש נקודות לעשור — עשרים נקודות ויותר על פני חיי אדם, באוכלוסיות שמאגר הגנים שלהן לא השתנה. בארגנטינה נרשמה עלייה של עשרים ושתיים נקודות במבחן אחד בין 1964 ל-1998. ומאז שנות התשעים תועדה בכמה מדינות מגמה הפוכה. המחקר החד ביותר בנושא בדק 817,611 מגויסים נורווגים שנולדו בין 1962 ל-1991: עלייה של כ-0.26 נקודות בשנה עד ילידי 1975 בערך, ואחר כך ירידה של כ-0.23 נקודות בשנה. והנקודה המכרעת: את כל התבנית הזאת — העלייה, נקודת המפנה והירידה — אפשר לשחזר **בתוך משפחות**.* כלומר בין אחים. זה שולל ישירות הסברים שמבוססים על הרכב האוכלוסייה.

מכאן אפשר לנסח את הנקודה שהיא הדבר החשוב ביותר בכל היחידה. אומדן תורשתיות **בתוך** קבוצה אינו אומר דבר על הסיבה להפרש **בין** קבוצות. זו אינה עמדה פוליטית אלא חשבון: תורשתיות בתוך קבוצה יכולה להיות גבוהה מאוד בזמן שהפרש בין שתי קבוצות נובע כולו מסביבה, ואין בכך שום סתירה. הפרשי ציונים בין קבוצות אוכלוסייה אכן נמדדים, ודוח הוועדה של 1996 כתב עליהם בפירוש: "בוודאי שאין תמיכה

בפרשנות גנטית. נכון להיום, איש אינו יודע מה מסביר את ההפרש הזה". העדכון המדעי מ־2012 הרחיק לכת יותר לכיוון הסביבה: ההפרש הצטמצם בכשליש סטיית תקן, וילדים שאומצו ממשפחות מעוטות אמצעים למשפחות מבוססות הראו עלייה של שתיים עשרה עד שמונה עשרה נקודות.

ולסיום, שתי אזהרות סימטריות. אחת: אל תכתבי שמנת משכל היא גורל. הכנסה — 0.23. שנה נוספת של בית ספר מזיזה את הציון בכמה נקודות. שתיים: אל תכתבי שהמבחנים "לא מודדים כלום" או שהם סתם מוטים תרבותית. מבחן מ־1932 ניבא מי חי בגיל שבעים ושש. ואפילו ההסבר הפופולרי להפרשים — "איום הסטריאוטיפ" — לא עמד בבדיקה של עצמו: במטא־אנליזה על 47 גדלי אפקט האפקט התכווץ מ־0.22 ל־0.07 ולא מובהק אחרי תיקון להטיית פרסום, והמחקרים המכוננים בתחום כללו כעשרה נבדקים לתא. גם הסבר נוח נבדק כאן כמו כל השאר.

למה זה ברשימה

כי זה הנושא שבו שתי הטעויות ההפוכות — שזה הכול ושזה כלום — נפוצות באותה מידה, ושתייהן ניתנות להפרכה במספרים.

שנה	2018
תחום	אישיות ואינטליגנציה
נושא	אינטליגנציה

להעמיק

אינטליגנציה: הידוע והלא ידוע

לקריאה

דוח הוועדה של האיגוד הפסיכולוגי האמריקני מ־1996, כולל הניסוח המפורש על תורשתיות ועל היעדר תמיכה בפרשנות גנטית להפרשים בין קבוצות. באנגלית.

<http://webspace.ship.edu/cgboer/iku.html>

להחכים לעד

לקריאה

הסבר בעברית על אפקט פלין ועל גודלו; שימי לב שהוא נכתב לפני שתועדה ההיפוך ומניח שהמגמה תימשך.

<https://davidson.org.il/read-experience/scientificamerican/לעד-להחכים-לעד/>

אינטליגנציה: ממצאים חדשים והתפתחויות תיאורטיות

מקורות וארכיון

העדכון המדעי לדוח: אימוץ למשפחה מבוססת מעלה את הציון בשתיים עשרה עד שמונה עשרה נקודות. באנגלית.

<https://library.scottbarrykaufman.com/uploads/2012/01/Nisbett-et-al.-2012.pdf>

מנת משכל בגיל אחת עשרה והישרדות עד גיל שבעים ושש

מקורות וארכיון

המעקב אחרי ילדי אברדין שנבחנו ב־1 ביוני 1932. באנגלית.

<https://pmc.ncbi.nlm.nih.gov/articles/PMC30556>

אינטליגנציה והצלחה חברתית־כלכלית

מקורות וארכיון

מטא־אנליזה של מחקרי אורך, עם המספר שמרסן את הציפיות: 0.23 להכנסה. באנגלית.

<https://gwern.net/doc/iq/ses/2007-strenze.pdf>

האם איום הסטריאוטיפ משפיע על ביצועי בנות

מקורות וארכיון

מטא־אנליזה שבה האפקט מתמוטט מ־0.22 ל־0.07 אחרי תיקון להטיית פרסום. באנגלית.

<http://icelab.psych.uw.edu.pl/wp-content/uploads/2018/10/florewicherts2015-Stereotype-threat.pdf>

ספרים

?Are We Getting Smarter

2012 · James R. Flynn

האיש שעל שמו נקרא האפקט מסביר מה הוא לדעתו כן אומר ומה לא — ולמה אין פירושו שאנשים נעשו חכמים יותר. לא אותר תרגום עברי.

Intelligence: All That Matters

2015 · Stuart Ritchie

הגנה מסודרת וקצרה על תוכנית המחקר, עם המגבלות שלה כתובות בגלוי. לא אותר תרגום עברי.

The Mismeasure of Man

1981 · Stephen Jay Gould

הביקורת הקלאסית על דטרמיניזם ביולוגי במדידת אינטליגנציה. כדאי לדעת שחלק מהניתוח ההיסטורי שבו עצמו שנוי במחלוקת — מה שמתאים בדיוק לרוח הפקולטה. לא אותר תרגום עברי.

מה באמת עובד בטיפול

2019 · שבעים שנות ראיות, ומספר אחד שמתכווץ

היחידה הזאת עוסקת בראיות ובמנגנונים בלבד. היא אינה ממליצה לאיש מה לעשות, במה לבחור או מתי לפנות; לשאלות האלה יש מקום, והוא אינו אתר לימוד.

הכול התחיל בהתקפה. ב־1952 פרסם הנס אייזנק סקירה של תשעה עשר מחקרים ויותר משבעת אלפים מקרים, והציג מספרים מביכים: שיפור בפסיכואנליזה 44 אחוז, בטיפול אקלקטי 64 אחוז — ובקבוצות שלא קיבלו טיפול כלל, 72 אחוז. שווה לדעת שהוא לא כתב "טיפול לא עובד"; הוא כתב שנראה מוקדם מדי להתעקש על הכללת הכשרה בטיפול בתוכנית הלימודים של פסיכולוגים קליניים. זו הייתה תקיפה על הכשרה, לא רק על יעילות.

התשובה הגיעה ב־1977, ובכלי חדש. סמית' וגלאס אספו 375 מחקרים מבוקרים, 833 גדלי אפקט, כעשרים וחמישה אלף נבדקים, וחישבו ממוצע: 0.68. בניסוח שלהם, המטופל הטיפוסי היה במצב טוב יותר מ־75 אחוז ממי שלא טופלו. הם גם מצאו, כבר אז, את הדבר שנקרא "פסק הדודו" על שם החיה מאליס בארץ הפלאות שהכריזה שכולם ניצחו וכולם יקבלו פרס: "מעט הבדלים חשובים ביעילות נמצאו בין סוגי טיפול שונים מאוד זה מזה".

היום אפשר לשים על פסק הדודו מספר. בסקירה על 409 ניסויים ו־52,702 מטופלים, ההפרש בין טיפול קוגניטיבי־התנהגותי לבין פסיכותרפיות אחרות הוא 0.06, כלומר לא מובהק ברוב הניתוחים; במונחים מעשיים, כדי שאדם אחד ייהנה מהמעבר בין הגישות צריך להעביר שישים ושלושה אנשים. יש גם צד שני לוויכוח: בהשוואות שבודקות את המדד הספציפי שהטיפול מכוון אליו נמצאים הבדלים קטנים אך אמיתיים, שנעלמים במדדים כלליים יותר.

ומה כן מסביר? הגורם היחיד שנמדד בעקביות הוא הברית הטיפולית — איכות הקשר בין המטפל למטופל. במטא־אנליזה של 295 מחקרים ויותר משלושים אלף מטופלים המתאם עם התוצאה הוא 0.278, כלומר כשמונה אחוזים מהשונות. המחברים עצמם מסייגים: אי אפשר להקצות אנשים באקראי לברית טובה או רעה, ולכן זהו קשר חזק ומתואם, לא הוכחה סיבתית. והמספר שמתגלגל ברשת, שהברית מסבירה שבעים אחוז מהתוצאה, פשוט אינו קיים בשום מקום.

עכשיו החלק שמחבר את היחידה לכל השאר. ב־2010 נבדקה הטיית הפרסום בספרות הזאת: 117 ניסויים, וגודל אפקט שיורד מ־0.67 ל־0.42 כשמשחזרים סטטיסטית כחמישים ואחד מחקרים שלא פורסמו. בבדיקה אחרת נמצא ש־58 אחוז מהתוצאות היו מובהקות בזמן שלפי העוצמה הסטטיסטית היו צפויות 49 אחוז — כלומר יש עודף תוצאות מוצלחות; ועוצמת המחקר החצינית בתחום היא 49 אחוז, מה שאומר שהמחקר הטיפוסי הוא הטלת מטבע.

ב־2019 הורץ החשבון כולו מחדש, במאמר שכותרתו "האם אייזנק צדק בכל זאת?". המסלול המלא, על 369 השוואות: 0.70 גולמי; 0.51 כשמוציאים מחקרים שמשווים לרשימת המתנה; 0.38 כשמסתכלים רק על מחקרים באיכות גבוהה; 0.31 אחרי תיקון להטיית פרסום. רק 23 אחוז מהניסויים עמדו בקריטריונים של סיכון הטיה נמוך. התשובה לשאלה שבכותרת היא "חלקית כן": האפקט אמיתי, והוא קטן בהרבה ממה שנמכר.

ולבסוף, ההודאה הכי חשובה בתחום. בסקירה מקצועית של ספרות המנגנונים נכתב במפורש שאף גורם — לא משותף ולא ספציפי — אינו עומד בקריטריונים של מנגנון פעולה מאומת אמפירית. כלומר: יש ראיות טובות שטיפול עוזר; אין עדיין הסבר מאומת ל**למה**. זה לא כישלון. זו פשוט התשובה נכון להיום, ויש בה יותר כבוד מאשר בהסבר יפה שאיש לא בדק.

למה זה ברשימה

כי זה התחום שבו הפער בין מה שהתחום יודע לבין מה שהוא מסוגל להסביר הוא הגדול ביותר — והוא מודה בזה בעצמו.

שנה	2019
תחום	המקצוע עצמו
נושא	ראיות בטיפול

להעמיק

אייזנק 1952: המאמר שהאשים את המקצוע כולו

לקריאה

הטקסט המקורי, קצר ובהיר, עם הטבלאות שמהן נולד כל הוויכוח. באנגלית.

<https://psychclassics.yorku.ca/Eysenck/psychotherapy.htm>

האם אייזנק צדק בכל זאת

לקריאה

החשבון המלא, שבו גודל האפקט יורד מ־0.70 ל־0.31 כשמנפים איכות והטיית פרסום. באנגלית.

<https://pmc.ncbi.nlm.nih.gov/articles/PMC6999031/>

קוקרן: טיפול התנהגותי מול שיטות אחרות

לקריאה

סיכום זהיר של עשרים וחמישה ניסויים, שמסכם שהשיטות יעילות במידה דומה — בראיות באיכות נמוכה עד בינונית. באנגלית.

https://www.cochrane.org/evidence/CD008696_behavioural-therapies-versus-other-psychological-therapies-depression

הדרכים לריפוי מחלות נפש

לקריאה

סקירה בעברית של המתח בין הגישה הביולוגית לפסיכולוגית ושל הדרך שבה נכנס לכאן תקן של רפואה מבוססת ראיות.

<https://davidson.org.il/read-experience/firefly/שפ/הדרכים-לריפוי-מחלות-נפש/>

הברית הטיפולית: המטא־אנליזה הגדולה

מקורות וארכיון

295 מחקרים, שלושים אלף מטופלים, מספר אחד — וההודאה שאי אפשר לרנדמז אותו. באנגלית.

<https://societyforpsychotherapy.org/wp-content/uploads/2018/10/Fluckiger-et-al-2018-Alliance-MA-Online.pdf>

ספרים

The Great Psychotherapy Debate

2015 · Bruce Wampold & Zac Imel

הניסוח הקנוני של מחנה "הגורמים המשותפים": שהברית והציפייה עושות את העבודה, לא הטכניקה. לא אותר תרגום עברי.

?What Works for Whom

2020 · Anthony Roth & Peter Fonagy

הצד הנגדי: מיפוי שיטתי של אילו גישות נבדקו מול מה, ובאיזו איכות ראיות. לא אותר תרגום עברי.

House of Cards

1994 · Robyn M. Dawes

הביקורת הקלאסית על הפער בין ביטחון קליני לבין ראיות — ומתחברת ישירות ליחידה על ניבוי מומחים. לא אותר תרגום עברי.

משבר השחזור ומה הוא שינה

2015 · איך תחום שלם גילה שהצינור שלו מייצר שטויות

בשנת 2011 פרסם דריל בם, פסיכולוג ותיק ומכובד, מאמר בכתב העת המוביל בפסיכולוגיה חברתית. תשעה ניסויים, כאלף נבדקים, שמונה מתוך תשעה מובהקים סטטיסטית. הממצא: אנשים מגיבים לגירויים **לפני** שהוצגו להם. על-יחשיות.

מה שהחריד את התחום לא היה הטענה אלא השיטות. הן היו רגילות לגמרי — בדיוק אותן שיטות שכולם השתמשו בהן שנים. ואם הצינור הזה מייצר הוכחה לניבוי העתיד, מה עוד הוא מייצר?

שלושה חוקרים ניסו לשחזר את אחד הניסויים בשלוש אוניברסיטאות, מאה וחמישים נבדקים. התוצאה: אפס מוחלט. את השחזור דחו שלושה כתבי עת מובילים, כולם בנימוק שזהו שחזור. באחד המקרים ההסתייגות הגיעה משופט שהתברר לימים כבם עצמו. המדיניות ההיא — "איננו מפרסמים שחזורים" — לא הייתה זדון; היא הייתה מקובלת. וזו נקודת התורפה המבנית. באותה שנה נסרקו 321,411 מאמרים בפסיכולוגיה מאז 1900, ונמצא ששחזורים הם 1.07 אחוז מהם — ומתוכם רק ארבעה עשר אחוז שחזורים ישירים. שחזור בידי אותם מחברים הצליח ב-91.7 אחוז מהמקרים; בידי מחברים אחרים — ב-64.6.

ובאותה שנה בדיוק פורסם המאמר שהסביר את המנגנון. סימנס, נלסון וסימונסון טבעו את המונח "דרגות חופש של החוקר": כל ההחלטות הלגיטימיות-לכאורה שחוקר מקבל בדרך — לאסוף עוד נבדקים? להוציא תצפיות חריגות? לפקח על משתנה? על איזה מדד לדווח? הם הראו בסימולציה שכל אחת לבדה מעלה את הסיכוי לתוצאה מובהקת-שגויה לשבעה עד שנים עשר אחוז, ושארבע מהן יחד מעלות אותו ל-60.7 אחוז — במקום החמישה שמובטחים. ואז הדגימו: עשרים סטודנטים הקשיבו לשיר של הביטלס או לשיר אחר, ולפי הדיווח, מי ששמע את הביטלס היה **צעיר בשנה וחצי**.* הטריק היחיד היה לפקח על גיל אביו של הנבדק. בלעדיו אין שום אפקט. הם לא זיפו דבר.

ב-2015 הגיע המספר שכולם מצטטים. מאתיים ושבעים חוקרים שחזרו מאה מחקרים משלושה כתבי עת מובילים: 97 אחוז מהמקוריים היו מובהקים, ו-36 אחוז מהשחזורים. גדלי האפקט התכווצו לכמחצית. שווה לדייק גם כאן, כי גם הנתון הזה מצוטט לא נכון: 47 אחוז מגדלי האפקט המקוריים נפלו בתוך רווח הסמך של השחזור, ו-68 אחוז נשארו מובהקים כשמאחדים את הנתונים. כישלון שחזור אינו הפרכה; לרוב פירושו "האפקט קטן בהרבה ממה שנטען, אם הוא קיים". והפילוח מסביר את כל הפקולטה: בפסיכולוגיה קונגיטיבית שוחזרו חמישים אחוז, ובפסיכולוגיה חברתית עשרים וחמישה.

היו גם פרויקטים שבדקו אפקטים קלאסיים בעשרות מעבדות בבת אחת. בראשון שוחזרו עשרה מתוך שלושה עשר, כולל עיגון ומסגור, ושניים נפלו לגמרי — שניהם פרימינג חברתי. בשני שוחזרו חמישה עשר מתוך עשרים ושמונה, וחציון גודל האפקט ירד מ-0.60 ל-0.15. והמסקנה החשובה: גודל האפקט היה קשור בעיקר

לאפקט שנבדק ולא למדגם או למקום — כלומר התירוץ "זה לא עבד כי הנבדקים שלכם שונים" ברובו לא החזיק.

וכאן אסור לעצור, כי החלק השני של הסיפור הוא שהתיקון עבד. נכנס לשימוש פורמט שבו כתב עת מאשר מחקר על בסיס השאלה והשיטה, לפני שיש תוצאות; מאתיים שבעים ושימונה כתבי עת אימצו אותו. וההשוואה מדברת בעד עצמה: בקרב מאמרים רגילים, 96 אחוז מדווחים על תוצאה חיובית; בקרב מאמרים שאושרו מראש, 44 אחוז. אותו דבר קרה ברפואה: בניסויים קליניים גדולים של מכון הלב האמריקני, לפני חובת הרישום המוקדם דיווחו 57 אחוז על תועלת מובהקת, ואחריה — שמונה אחוז.

ושתי אזהרות לסיום. אחת: זו אינה בעיה של פסיכולוגיה. חברת תרופות ניסתה לשחזר חמישים ושלושה מאמרי דגל בביולוגיה של הסרטן ואישרה שישה; בסקר של כתב עת מדעי על 1,576 חוקרים, מעל שבעים אחוז נכשלו בשחזור ניסוי של משהו אחר ומעל חמישים אחוז נכשלו בשחזור של עצמם. פסיכולוגיה לא נכשלה יותר מאחרים — היא **מדדה** יותר מאחרים. ושתיים, לצד השני: ב־2023 פורסם מאמר שהראה שאפשר להגיע לשחזוריות גבוהה, וב־2024 הוא נמשך, בין השאר בגלל הצהרה שגויה על רישום מוקדם. גם זה, כשחושבים על זה, ראייה שהמנגנון עובד.

למה זה ברשימה

כי כל תשע עשרה היחידות שלפניה הן דוגמאות, וזו התשובה לשאלה מה בעצם קרה כאן.

שנה	2015
תחום	המקצוע עצמו
נושא	משבר השחזור

להעמיק

שלושה ניסיונות שחזור לבם

לקריאה

השחזור שאף כתב עת מוביל לא הסכים לפרסם, ושהתוצאה בו אפס. באנגלית.

<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0033423>

בם הוכיח שיש עלי-חוש, כלומר שהמדע שבור

לקריאה

פרופיל שבו בם עצמו מתאר בלי התנצלות איך באמת עבד — וזה תיאור של פרקטיקה נורמלית. באנגלית.

<https://slate.com/health-and-science/2017/06/daryl-bem-proved-esp-is-real-showed-science-is-broken.html>

האם אפשר ללמוד ממחקרים פסיכולוגיים

לקריאה

הצגה עברית נקייה של פרויקט השחזור ושל פרשת הזיוף ההולנדית, שמסיימת בשאלה הנכונה.

<https://davidson.org.il/read-experience/sciencenews/פסיכולוגיים-ממחקרים-ללמוד-האם-אפשר-ללמוד>

תשעים ושבעה ניסויים הצליחו במקור, שלושים ושישה שרדו

לקריאה

כתבה עברית שמתרגמת את המספרים למה שנשאר מאפקטים מפורסמים שהופיעו ביחידות הקודמות.

<https://www.bizportal.co.il/science/news/article/20037206>

תשעים ושישה אחוז מול ארבעים וארבעה

לקריאה

ההשוואה בין מאמרים רגילים למאמרים שאושרו מראש — הראיה הטובה ביותר שהתיקון עובד. באנגלית.

<https://journals.sagepub.com/doi/10.1177/25152459211007467>

פסיכולוגיה של חיוביישגוי, 2011

מקורות וארכיון

המאמר שהמציא את "דרגות החופש של החוקר" והוכיח אותן בכך שהראה שהביטלס מצעירים אנשים. באנגלית.

<https://www2.psych.ubc.ca/~schaller/528Readings/SimmonsNelsonSimonsohn2011.pdf>

ספרים

Science Fictions

2020 · Stuart Ritchie

נכתב בידי אחד משלושת החוקרים שהשחזור שלהם נדחה מכל כתב עת, כלומר בידי עד ראייה. לא אותר תרגום עברי.

The Seven Deadly Sins of Psychology

2017 · Chris Chambers

מאת מי שהמציא בפועל את פורמט האישור-מראש שמופיע בסוף היחידה. לא אותר תרגום עברי.

Rigor Mortis

2017 · Richard Harris

מרחיב את אותה בעיה לביורפואה, ומועיל בדיוק לשאלה אם זו בעיה של פסיכולוגיה בלבד. לא אותר תרגום עברי.

קולופון

פסיכולוגיה — חוברת קריאה מתוך האוניברסיטה, ספריית הלמידה האישית של מושקא.

20 יחידות · 110 קישורים מאומתים · 60 ספרים. הגרסה המלאה, עם חיפוש וכרטיסיות חזרה, נמצאת ב-
uni.zikeet.co.il

נבנתה מאותם קובצי תוכן שמהם נבנה האתר, כך ששתי הגרסאות אינן יכולות להיפרד זו מזו.