

מה אומר ה-IQ?

מהם בדיוק מבחני IQ, ומה מקורם? כדי לדון דיון אינטליגנטי על מבחני אינטליגנציה, עלינו להשיב על שאלות אלה. ראשית לכול, קיימות שתי מסורות במחקר האינטליגנציה.

פרנסיס גלטון: מכשיר הדחיקה נרחק החוצה

אחד מן הספרים בעלי ההשפעה הרבה ביותר בכל הזמנים היה ספרו של צ'רלס דארווין **מוצא המינים** (Darwin, 1859). דארווין טען בספר זה שניתן להתחקות אחר האבולוציה של המינים והתפתחות האדם באמצעות התהליך האבולוציוני של הברירה הטבעית. הספר השפיע מאוד על מחקרים רבים, שאחד מהם הוא חקר האינטליגנציה האנושית והתפתחותה. דארווין הציע שיכולותיו של האדם המשכיות במובן מסוים לאלה של בעלי החיים. מהו, אם כן, הקשר בין יכולותיהם של בעלי חיים לבין אלה של בני האדם?

בן דודו של דארווין, סיר פרנסיס גלטון (Galton), היה כנראה הראשון שחקר את משמעותו של הספר לגבי האינטליגנציה האנושית. גלטון טען שקיימות שתי תכונות כלליות המבחינות בין בני אדם אינטליגנטים יותר או פחות. הראשונה היא אנרגיה, או היכולת לעבוד. גלטון סבר שבעלי כשרים אינטלקטואליים בתחומים מגוונים מאופיינים על ידי רמות גבוהות של אנרגיה. התכונה השנייה היא רגישות. לפי גלטון, ככל שאנו אינטליגנטים יותר, כך אנו רגישים יותר לגירויים שמסביבנו. עד כאן דבריו נשמעים מדעיים מאוד; אך מאמציו הראשונים של גלטון בחקר האינטליגנציה לקו במיזוג של מדע עם אמונות טפלות שרודף את התחום עד עצם היום הזה. לדוגמה: "יכולת ההבחנה של אידיויטים נמוכה

במיוחד. הם מבחינים בקושי בין חוס לקור, ותחושת הכאב שלהם קהה כל כך עד שאידיויטים ברמה הנמוכה אינם חשים לכאורה כאב כלל. בחייהם חסרי הרגש, הכאב שניתן לעורר בהם מתקבל בשמחת הפתעה".¹ במשך שבע שנים – משנת 1884 ועד 1890 – ניהל גלטון שירות במוזאון של דרום קנסינגטון בלונדון שבו יכול כל אחד, תמורת סכום פעוט, לבחון את האינטליגנציה שלו. הבעיה היחידה הייתה שהמבחנים שהשתמש בהם היו תערובת מוזרה שמדדה דברים שונים, אך על כל פנים לא את האינטליגנציה. גלטון המציא בין היתר משרוקית שניתן היה לקבוע באמצעותה את הטון הגבוה ביותר שהנבחן מסוגל לשמוע. זוהי בדיקה מעניינת של רגישות שמיעתית, אך אין זה מבחן אינטליגנציה. אם אכן המבחן אמור למדוד אינטליגנציה, הרי שכל חתול חכם בהרבה מכל אחד מאתנו.

במבחן שני השתמש גלטון במספר אריזות של תרמילי רובה ציד מקוריים או ממולאים בצמר או בצמר גפן. כלפי חוץ נראו האריזות זהות, ורק הכובד הבחין ביניהן. על הנבחן היה להבדיל בין האריזות הקלות לכבדות. האם זהו מבחן אינטליגנציה או מבחן רגישות למשקל? הוא הדין במבחן הרגישות לריח ורדים. זהו מבחן שיש להימנע ממנו אם אתה מצונן, או סובל מאלרגיה לוורדים.

ניתן היה לשער שגלטון זכה ללעג ולקלס; אך ההפך הוא הנכון. הוא לא קיבל את התואר "סיר" בשל היותו בדיחה מהלכת. עבודתו נתקבלה ברצינות רבה, ועד כמה שידוע לנו, הוא הרוויח הון רב מפעילות זו. לא רק מבקר המוזאון שבא לקבל הערכה עצמית מדעית קיבל את דברי גלטון ברצינות. פסיכולוג מפורסם בשם קטל (Cattell) התרשם מרעיונותיו עד כדי כך שהביא אותם אתו לאמריקה.

קטל פיתח בשנת 1890 מבחן משלו, הדומה בעיקרו לזה של גלטון. במבחן דחיקה באמצעות מד-כוח, היה על הנבחן לדחוק את המכשיר בכל כוחו, וכוח הדחיקה שימש כאחד ממדדי האינטליגנציה. מבחן אחר בדק את מידת הלחץ שיש להפעיל על הנבחן עד שיחוש כאב (זאת הייתה הזדמנות פז לאלה הנפגעים בקלות – הם זכו לציון גבוה יותר בסולם האינטליגנציה).

כפי שניתן לשער, בעיה אחת הייתה משותפת לכל המדדים הללו: הם לא פעלו. לבסוף, תלמיד של קטל בשם ויסלר (Wissler) מצא שתוצאות

המבחנים של קטל לא היו קשורות זו לזו, וגם לא לציוני הנבחנים באוניברסיטה בה עבדו קטל וויסלר – קולומביה. גם עבודתו של ויסלר לא הייתה מחקר מדעי למופת, אך היה בה די כדי להוכיח ששיטתו של קטל היא כישלון חרוץ.

למרבה הצער, בזה לא תם ונשלם. כפי שקבע סנטיאנה, מי שאינו לומד מן ההיסטוריה סופו שהוא חוזר עליה. כיום ישנה קבוצה של נאו-גלטונים שהפихו רוח חיים בעבודתם של גלטון וקטל. הם משתמשים במדדים כגון זמן תגובה פשוט (באיזו מהירות אתה לוחץ על כפתור לאחר ראיית הבזק של אור) וזמן זיהוי קו (באיזו מהירות תזהה איזה משני קווים ארוך יותר), כדי למדוד אינטליגנציה. בעסק הזה, אסור לטעות ולחשוב שרעיון שנקבר אכן נעלם לנצח – הוא עלול לחזור ולרדוף אותנו. ברעיונותיו של גלטון, שעוצבו בהשראת תורת האבולוציה של דארווין, היה אכן היגיון אבולוציוני כלשהו. חיות בר בעלות חסך באחד החושים אינן זוכות לחיים ארוכים, ובעלות החושים החדים מצליחות יותר. ייתכן שלפני שנים רבות היה יתרון בבררה הטבעית לאלה מבין אבותינו שהיו בעלי חושים חדים במיוחד. עם זאת, מה שמוביל ליתרון סלקטיבי בתקופה מסוימת אינו מועיל במיוחד בתקופה אחרת. ידועה הדוגמה של עשים בהירים וכהים החיים בעיר, בעלי יתרונות המשתנים בהתאם לרמת זיהום האוויר: הכהים אינם גלויים לעין ברמה גבוהה של זיהום אוויר, בעוד שהבהירים גלויים יותר; היתרון מתהפך כאשר הזיהום פוחת. בתקופתנו, יכולת חושית גבוהה אינה גורמת ליתרון בהתרבות או להישרדות טובה יותר. יש תאוריות המופיעות טרם זמן; התיאוריה של גלטון איחרה בכמה אלפי שנים.

אלפרד בינה: הערצת הכשרים האקדמיים

בשנת 1904, הקים השר לענייני חינוך בפריז ועדה שנועדה לחפש דרכים להבחנה בין ילדים "מפגרים" באמת לבין אלה שאינם מצליחים בבית הספר מסיבות אחרות. תפקיד הוועדה היה להבטיח שילדים יופנו לכיתות למפגרים רק במקרה ש"אינם מסוגלים להפיק תועלת באופן ממוצע מן

ההוראה הניתנת בבתי הספר הרגילים".² ד"ר אלפרד בינה (Binet) ועמיתו תאודור סימון (Simon) פיתחו מבחנים שיענו על צורך זה.

יש לציין שעבודתו של בינה נבעה מן השאיפה לעזור לילדים ולהגן עליהם, ולא להעניש אותם. מורים שבכיתתם היו תלמידים שהפריעו, מצאו דרך שנראתה להם כהקלה עצומה: הם יכלו להמליץ להעבירם לכיתת מפגרים. המורים לא היו בהכרח מנוולים ששאפו רק להקל על חייהם (אם כי מחשבה זאת בוודאי חלפה במוחם). לא הייתה אבחנה ברורה בין ילדים בעלי בעיות התנהגות לבין המפגרים שכלית, וילדים בעלי בעיות התנהגות טופלו כאילו היו מפגרים.

תפיסתם של בינה וסימון על אינטליגנציה ועל הדרכים למדוד אותה שונה מאוד מזו של גלטון וקטל. בינה וסימון כינו את מבחניהם של הללו "בזבוז זמן", והתייחסו לליבת האינטליגנציה כ"שיפוט, הנקרא גם שכל ישר, שכל מעשי, יוזמה, יכולת הסתגלות לנסיבות. לשפוט היטב, להבין אל נכון, לשקול כראוי – אלה הן הפעילויות העיקריות של האינטליגנציה".³

נודה על האמת: תפיסתו של בינה סבירות בהרבה מאלה של גלטון. בינה הביא את הקלן קלר כדוגמה לאדם שציוניו במבחני רגישות של שמיעה וראייה הם אפסיים, ועם זאת רמת האינטליגנציה שלו מוכרת כיוצאת מן הכלל. בינה פיתח מבחנים שבהם גם נכה יוכל להצליח. הוא תפס את האינטליגנציה כתלויה בשיפוט מנטלי ולא בחדות החושים.

מרביתנו מכירים את בינה רק באמצעות המבחן שלו; אך הוא פיתח גם תיאוריה טובה על האינטליגנציה. הוא טען שמחשבה אינטליגנטית בנויה משלושה יסודות מובחנים, אותם כינה **כיוון, הסתגלות וביקורת**. **כיוון** הוא הידיעה מה נדרש לעשות וכיצד לעשות זאת; **הסתגלות** היא פיתוח אסטרטגיה לביצוע מטלה, מעקב אחריה והתאמתה תוך כדי פעולה; **ביקורת** היא היכולת לבקר את מחשבותיך ואת מעשיך. אם ברצונך למשל לקנות מכונת חדשה, תשתמש בכיוון כדי לוודא מה אתה חייב לדעת על מכונות וכיצד ליישם מידע זה לשם בחירה הלכה למעשה. הסתגלות תהיה מעורבת בחיפוש, בקבלת החלטות ואולי גם בשינוי אסטרטגיה במשך החיפוש – למשל, החלטה שלא לרכוש טלוויזיה לנוסעי המושב האחורי. ביקורת היא הערכה על דרך קבלת ההחלטות,

והצגת השאלה האם אתה נוקט באסטרטגיה נאותה להחלטה על דגם המכונית ועל המקום שבו תרכוש אותה. רעיונותיו של בינה, כמו אלה של גלטון, יובאו לארצות הברית – במקרה זה לקליפורניה ולא לניו יורק. לואיס טרמן (Terman), פרופסור לפסיכולוגיה באוניברסיטת סטנפורד, פיתח מבחן אמריקני המבוסס על התיאוריה ועל המבחנים של בינה. מבחן זה היה לציון דרך בהיסטוריה. "מבחן סטנפורד-בינה" מהווה עד היום תחרות רצינית בזירת מבחני האינטליגנציה.

רעיונותיו של בינה היו הגיוניים בהקשר שבו הוצעו: ניבוי הצלחה אקדמית. למרבה הצער, כמו רעיונות רבים שהצליחו, נעשה גם בו שימוש הרבה מעבר לתחום שבו הוא פועל במיטבו. מבחנים של כשרים אקדמיים הועברו לשימוש גם בתחומים שאינם אקדמיים. זאת ועוד, יש למבחנים מגבלות המטילות ספק בתוצאותיהם גם בהקשרים אקדמיים רבים.

מבחנים המבוססים על התיאוריה של בינה

מהו סוג השאלות המופיעות במבחן IQ? רבים מאתנו שמעו על מבחני IQ, ורבים התנסו בהם במהלך חייהם, אך קשה לעתים לזכור את פרטיהם. למעשה, רבים דנים במבחני IQ ללא ידיעה ברורה מה יש בהם ומה פגום בהם. מלל רב מבוסס על אמונות פוליטיות או חברתיות ולא על עובדות. אין פסול בהסקת מסקנות פוליטיות או חברתיות מן המבחנים – אני עצמי עושה כן; אך ראשית עלינו להבין את המבחנים, ולדעת מה נכון ומה לא נכון בהם. אתאר כאן בפירוט מבחן אחד מן השניים הנפוצים ביותר – מבחן סטנפורד-בינה, שצמח מן המבחנים המקוריים של בינה.

סולם האינטליגנציה של סטנפורד-בינה, מהדורה רביעית (The SB IV – Stanford-Binet Intelligence Scale, Fourth Edition) הוא האחרון בסדרה של סולמות שתחילתם בשנת 1905.⁴ הנוסח המתוקן הראשון (כלומר, המהדורה השנייה) של המבחן הופיע בשנת 1937, והמהדורה השלישית – ב-1960. ניתן להעביר את המבחן לילדים מגיל שנתיים ולמבוגרים בכל גיל שהוא, אף על פי שהרצות הניסיון של המבחן (התקנון) כללו נבדקים עד לגיל 23 בלבד.

מה יש במבחן? כיצד נראה מבחן סטנפורד-בינה? ישנם בסך הכול חמישה עשר תת-מבחנים, שרק שישה מהם משמשים לכל טווח הגילים שהמבחן מכסה אותו. התת-מבחנים מחולקים לארבע קטגוריות: חשיבה מילולית, חשיבה כמותית, חשיבה פיגורטיבית/מופשטת וזיכרון לטווח קצר. נתבונן במספר דוגמאות מתוך חמישה עשר המבחנים:

1. **אוצר מילים.** הנבחנים מתבקשים לציין פירושן של מילים. ברמות הנמוכות, המילים מוצגות באמצעות תמונות; לאחר מכן הן מוצגות בכתב (לדוגמה, המילה "יומרני"). מבחן אוצר המילים נערך לכל רמות הגילים. שאלות הבוחנות אוצר מילים מופיעות במבחני אינטליגנציה רבים ובמבחנים דומים אחרים, כגון SAT. הכללת יסודות אלה מבהירה לנו את ספקותיהם של פסיכולוגים ומורים בדבר תיוגם כמבחני אינטליגנציה.

ראשית, יש להבין שההבחנה רבת-השנים בין מבחני אינטליגנציה מצד אחד לבין מבחני הישגים מצד שני, היא בעיקרה מיתית. מבחני אינטליגנציה מודדים ככלל את ההישגים שהנבחן רכש לכאורה בשנים שקדמו למבחן. ברור שמבחני האינטליגנציה מודדים הישגים – מהו אוצר מילים אם לא הישגים? אף אחד לא יטען – אף לא הדוגל הנלהב ביותר בנטיביזם מבין התאורטיקנים – שאנו נולדים כשאוצר מילים מאוחסן בנוחיות במוחנו. נדרשת אינטליגנציה גם כדי לעמוד במבחני הישגים – ללמוד את החומר המופיע במבחן.

שנית, כאשר בוחנים אוצר מילים, אפשר לוותר מראש על כל אשליה בדבר הוגנות תרבותית. מבחן של אוצר מילים באנגלית לא ימדוד אצל ילד ששפת אמו ספרדית, וייטנאמית או יפנית את מה שימדוד אצל ילד ששפת אמו אנגלית. מספר המילים הכולל שבו בקיא ילד דו-לשוני עשוי להיות גדול מזה שיודע ילד המדבר אנגלית בלבד; אך המבחן מתגמל רק על המילים באנגלית, ומעניק בכך יתרון לילד ששפת אמו אנגלית.

למרבה הפלא, אלה המעוניינים להשתמש במבחני אינטליגנציה מחוץ לארצות שבעבורה הם פותחו (ארצות הברית היא היצרנית והצרפנית הגדולה ביותר של מבחנים אלה), מתרגמים אותם לשפת המקום. הם מניחים בכך שניתן לתרגם את המילים במדויק; שהמילים המתורגמות שוות במידת הקושי שלהן ובחשיבותן גם בשפה האחרת;

ושמתן הגדרות למילים מופשטות היא פעולה מוכרת בתרבויות אחרות כמו בתרבות האמריקנית. אף לא אחת מהנחות אלה נכונה בהכרח, אך עובדה זאת אינה מונעת הפקת תרגומים חסרי דעה של מבחנים אמריקנים או אחרים.

שלישית, הדרך שבה נבדק אוצר המילים היא דוגמה לנושא שיחזור ויופיע במהלכו של ספר זה: מבחני אינטליגנציה מודדים אינטליגנציה אקדמית, נטולת הקשר. לאחרונה ביקרתי בכיתה שבה בחן המורה את אוצר המילים של תלמידי תיכון. כל תלמיד התבקש לתת למילה הגדרה מופשטת ולהשתמש בה במשפט.

ההגדרה המופשטת של תלמיד אחד הייתה טובה יותר מזו של רעהו, אך השימוש בה בתוך ההקשר לא היה מדויק. תלמיד אחר לא היה מסוגל להציע הגדרה מופשטת, אך השתמש במילה בתוך ההקשר בצורה נאותה. תלמיד זה יחשב כאינטליגנטי יותר בפעילות היומיומית ואפילו בכיתה, בעוד שהתלמיד הראשון יקבל ציון טוב יותר במבחן ה-IQ.

אחרון חביב, ראו מה עוללו מבחני אוצר המילים למערכת החינוך שלנו. מדוע שינו התלמידים בבית הספר התיכון רשימות של מילים? זאת אינה דרך נאותה ללימוד קריאה או כתיבה, דיבור או הקשבה. הדרך הטבעית ללימוד מילים היא בהקשר – באמצעות קריאה והקשבה, למשל. בסופו של דבר, מה שחשוב בחיים הוא לדעת לקרוא, לכתוב, להקשיב ולדבר היטב, ולא לפלוט הגדרות. שינון בעל פה של אוצר מילים לא זו בלבד שאינו טבעי, אלא אינו מוביל לזכירת אותן מילים לטווח ארוך. הן נשכחות עד מהרה כי לא נלמדו בהקשר משמעותי כלשהו, כמו גם מרבית התכנים של השיעורים שאנו משננים בבית הספר. המורים מכריחים מיליוני תלמידים לבזבז את זמנם בשינון מילים שלא ידעו כיצד להשתמש בהן ובוודאי ישכחו אותן במהרה, בשל העובדה שמבחנים של כישרון לימודי מוערכים בבתי הספר. הוראת המורים מכוונת למבחנים, אך אלה אינם מודדים אותם דברים שילדים או הבוגרים צריכים לדעת.

2. **הבנה.** על הפרט להוכיח הבנה של נורמות חברתיות ותרבותיות – למשל, על ידי הסבר מדוע אנשים לווים לפעמים כסף, או מדוע אנשים מצביעים. במבט ראשון, נדמה שמבחן כזה אכן מודד אינטליגנציה המתייחסת לעולם האמיתי – כושר שאנו כמובן מעוניינים בו. האם הוא

אכן מודד הבנה של העולם האמיתי או רק קריקטורה של העולם האמיתי – סיפור שאנו רוצים לספר לעצמנו על החברה שבה היינו רוצים לחיות? מדוע אנשים לווים כסף? לכאורה, כי הם זקוקים לממון לרכישת משהו שאינם יכולים לשלם את תמורתו המלאה ברגע זה. אך מדוע אנשים באמת לווים כסף? לפעמים, כמו ברכישת בית, לצורך הקלות במס; לפעמים, מכיוון שהם חושקים במותרות – מכונית, סירה או דירת פאר – שלא זאת בלבד שאינם יכולים לרוכשם, אלא אינם צריכים לרוכשם; לפעמים, כדי שיהיו מצויים בכיסם מזומנים שאין להם כל כוונה להחזירם; ולפעמים, לרכישת סמים שיביאו בסופו של דבר למותם. לפעמים, כמובן, הם לווים כסף בשל הסיבות הניתנות במפתח התשובות לשאלות המבחן. מדוע אנשים מצביעים? במדינה שבה 99.6 אחוז מן האזרחים מצביעים לרודן השליט (השיעור הבלתי מפתיע של ההצבעה האחרונה לסדאם חוסיין), הרי זאת משום שהם מחויבים בכפייה להצביע למועמד המוצג מטעם הממשלה; לפעמים הם רוצים להביע אי נחת מכל המועמדים המוצעים (כמו במקרה של אזרח אמריקני הכותב שם של מועמד המועדף על ידו במקום להטיל לקלפי אחד מן הפתקים המוכנים), ולפעמים, משום שקיבלו תשלום כדי להצביע כך. מובן שלעתים הצבעה זאת משקפת את עמדותיהם הפוליטיות – התשובה הרצויה במבחן. אך קורה גם שלשאלה זאת אין כל משמעות, מכיוון שבמדינות רבות בעולם אין מתקיימות בחירות כלל, והאזרחים לא יצביעו לעולם.

מבחן הבנה בוחן בעיקר את מה שהנבחן יודע שהבוחן רוצה לשמוע, ולא את מה שהנבחן יודע. אם ברצונו למדוד הבנה של משחק המבחנים, הרי שמבחן כזה הוא מדד נאות. הוא עשוי להיות גם בעל מתאם להבנת אותו משחק בבית הספר. אין זה מדד של אינטליגנציה מצליחה, במיוחד לגבי אנשים שלא התחנכו להבנת משחק המבחנים. דווקא מבחן זה, הנראה תלוי תרבות כל כך, אינו קשור כלל לתרבות שבה אנשים חיים. הוא קשור לפנטזיה מסוגם של דן, דנה ורמה בדבר מה שהתרבות רוצה להאמין על עצמה.

3. **אבסורדיזם.** במבחן זה מוצגות בפני הנבחן תמונות המכילות פרטים לא סבירים, ועל הנבחן לזהות אותם. ייתכן שיצטרך להצביע על כך ששחקני הוקי קרח לא מחליקים על פני אגם שלתוכו מזנקים שחיינים

בבגדי ים. מה יודע ילד בארץ טרופית על הוקי קרח? האם אמריקני היה מזהה פרטים לא סבירים בתמונה המציגה משחק כדורגל אנגלי או קריקט, שאפילו אינם משחקים טרופיים? ומה בדבר משחק ה"ג'ו"? כדי לזהות אבסורדים יש להכיר את התוכן המוצג בתמונה, ולדעת מה עשוי להיות בלתי הולם בה.

4. **קשרים מילוליים.** כאן על הנבחן להשיב מה משותף לשלוש המילים הראשונות בקבוצה, שאינו משותף למילה הרביעית. מה משותף, למשל, לתפוח, בננה ותפוז, ולא לספל? – מטלה פשוטה למדי, אך קלה יותר לביצוע כאשר אתה בא ממקום שבו כל הפרות האלה מצויים ומוכרים לך. האם ידוע לילד אמריקני מהי גויאבה? ישנן ארצות שם כל ילד יודע מהי. לעתים קרובות מבחנים שאמורים לבחון חשיבה מילולית בוחנים למעשה אוצר מילים. שני ילדי נבחנו ב-Preliminary Scholastic Assessment Test, ושניהם העירו אותה הערה: האנלוגיות המילוליות האמורות לבחון חשיבה מילולית, היו קשות לא בגלל שהם לא יכלו לזהות את האנלוגיות, אלא בגלל שלא הכירו את כל המילים.

כך קורה לעתים שמבחן האמור למדוד דבר אחד (חשיבה מילולית) מודד דבר אחר (במקרה זה, אוצר מילים). לא ניתן לפתור שאלה באנלוגיות מבלי להבין את המילים.

הדבר המזיק במיוחד במבחנים אלה הוא שהם גורמים למפרשי הציונים להגיע למסקנה המוטעית, שמי שאינו מבין את פירושן של מילים נדירות לוקה בחסר בכושר החשיבה. לדעתי, טוב היה אילו כולם ידעו את פירושן של מילים נדירות. מובן שמי שבדעתו לפתח קריירה ספרותית, ראוי שיכיר מילים אלה. עם זאת, אני סבור שרבים יחיו את חייהם היטב, ואולי אף בהצלחה רבה, ללא הכרת מילים כאלה.

5-6. **ניתוח דגם והעתקה.** במבחנים של ניתוח דגם, הנבחן משחזר בעזרת קוביות דגמים דו-ממדיים בשחור לבן, ליצירת צורות גאומטריות. בהעתקה, על הנבחן לשחזר ציורי קו גאומטריים.

שני תת-מבחנים אלה מדגימים בבירור רב את מגבלת התפיסה המקובלת של האינטליגנציה. המונח "העתקה" מדבר בעד עצמו. ביסודו של דבר, הנבחן מתוגמל על העתקת עבודתו של אדם אחר. איזו מין תפיסה של אינטליגנציה היא זאת? הדבר דומה למתן ציון חיובי על יכולת

אמנותית מעולה לאנשים שיושבים במוזאון ומבלים את זמנם בהעתקת עבודותיהם של גדולי הציירים. בו בזמן, אין תימה שמבחנים כגון סטנפורד-בינה מנבאים הצלחה בבית הספר, מאחר שנדרשים בו פעמים רבות העתקה ושחזור של דברי המורה. אם אתה מבצע זאת היטב, אתה זוכה בציון "מעולה". אם אינך עושה זאת כראוי – אולי משום שאין לך עניין במטלה מכנית כזאת – האינטליגנציה שלך מוטלת בספק. אני חש לעתים שאחד היתרונות בהיותי מבוגר ולא ילד הוא בכך שאין מצפים ממני למלא ספרי צביעה בצבעים ה"נכונים", תוך הקפדה להישאר במסגרת הקווים. אם אנו מאמינים במבחן זה, הרי שספרי צביעה הם מדדי אינטליגנציה מצוינים.

7. **מטריצות.** בתת-מבחן זה מוצגות לנבחן מטריצות צורניות שחסר בהן חלק אחד, ועליו לבחור מבין מספר חלופות את זאת שתתאים לחלל החסר. למטריצות היסטוריה ארוכה בעסקי מבחני האינטליגנציה. ישנם המחשיבים את ה-Raven Progressive Matrices, המכיל בעיות בתחום של מטריצות צורניות בלבד, כאחת מצורות המדידה הטהורות ביותר של אינטליגנציה כללית. מבחן זה נחשב לכליל השלמות של הוגנות תרבותית, מכיוון שאין בו מילים והוא נראה לתומכיו כעומד מעל ומעבר לתוכן תרבותי. ייתכן שיש בו תכונות טובות רבות, אך הוגנות תרבותית אינה אחת מהן.

ראינו כמה סוגי פריטים שעשויים להופיע במבחני אינטליגנציה הניתנים ליחידים. נבחן עתה סוגיות נוספות הדנות בדרכים לחישוב הציונים ולפירוש התוצאות.

במדגם התקנון של סטנפורד-בינה – מדגם המשמש להפקת מערכת של המרה בין מספר התשובות הנכונות לבין ה-IQ – היה טווח של 22 שנה בין הנבחן הצעיר ביותר לבין זה המבוגר ביותר, וזהו מרווח ניכר ביותר. מדגם תקנון זה משמש כקבוצת השוואה לקביעת ציוניהם של הנבחנים. כמה אנשים נבחנו, לדעתכם, לצורך קביעת ההמרה בין מספר הפריטים שניתנה להם תשובה נכונה (מה שמכונה הציון הגס) לבין הציון הממשי במבחן (מה שמכונה מנת המשכל, שעליה נרחיב את הדיבור להלן)? מספר התושבים בארצות הברית כיום הוא כ-250 מיליון. האם נראה לכם מדגם בן 500 אלף, או אולי 50 אלף? מה דעתכם על 5,000? המספר המדויק היה אכן

5,013 – לא הרבה בהתייחס להחלטות המתבססות על הדיוק של תוצאות אלה ומשמעותיות כל כך לחיי בני אדם. יש לזכור גם ש-5,000 אלה מתפרסים על פני 22 שנות חיים, כך שמדובר בפחות מ-250 לכל שנת חיים!

מרווח של 22 שנים מרשים אמנם מנקודת מבט אחת, אך אינו מרשים כלל מנקודת מבט אחרת. מרווח זה מדגיש את העובדה שהמבחנים מודדים כשרים אקדמיים. המדגם תואם בדיוק את מספר השנים שתושבי ארצות הברית מבלים בלימודים. אך בני האדם ממשיכים בדרכם לאחר גיל 22 בערך, ומצטרפים לעולם העבודה. הכישורים הגורמים להצלחה משתנים, אך המבחנים אינם משתנים. אם כן, לפנינו מבחן שאינו רק בנוי, אלא אף מתוקן לפי מדגם מוטה-גיל של האוכלוסייה. מציאות זאת נחשפת כאשר אדם מבוגר מחליט לחזור ללימודים, ומוצא את עצמו חייב לעמוד במבחן המייצג את המציאות של מתבגר ולא של מבוגר. הוא מבזבז את זמנו בלימוד גאומטריה וטכניקות אחרות שלא השתמש בהן מזה שנים רבות, ושאלו כל סיכוי שישתמש בהן שנית אלא במבחנים.

אף על פי שנעשה מאמץ שהמדגם ייצג את גיל האוכלוסייה (מוטת הגיל) של ארצות הברית, לא הוכתר מאמץ זה בהצלחה. הקבוצה הזוכה תמיד לדגימת יתר במחקרים פסיכולוגיים, במיוחד במחקרי אינטליגנציה, היא זאת של בעלי רקע חברתי-כלכלי גבוה. לפיכך נעשו תיקונים סטטיסטיים כדי לפצות על דגימת יתר זאת, אך אלה הם ניחושים בלבד. זהו אם כן עוד הליך שאינו נוסך ביטחון בדיוקם של הציונים הסופיים.

השוואה עם מבחנים אחרים, כולל נוסחים קודמים של סטנפורד-בינה, מראה שאם כי הציונים בכללם הם בני השוואה, הציונים של המבחנים והמפגרים שכלית נוטים להיות נמוכים במקצת מציוניהם בנוסחים הקודמים של המבחן ובמבחן המתחרה – מבחן וכסלר. בהתחשב בעובדה שמקבלי ציונים נמוכים מאובחנים כ"מפגרים שכלית", וציונים גבוהים מקנים למקבליהם תווית של "מחוננים", הרי שאי היכולת השיטתית לערוך השוואה של התוצאות עלולה לפגום בביטחוננו במבחן שנודעת לו משמעות רצינית לחיים.

קיימות במבחן גם בעיות אחרות – למשל, סוגים שונים של תת-מבחנים (מטלות) ניתנים לקבוצות גיל שונות, על כן לא ברור אם אפשר להשוות ציון בגיל אחד עם ציון בגיל אחר. זאת ועוד, טווח הציונים

האפשריים שונה בגילים שונים, בתת-מבחנים שונים ובסדרות של תת-מבחנים – עובדה שמקשה על השוואת הציונים לא רק בגילים שונים אלא גם בסוגים שונים של כשרים (קביעת דגם של נקודות העוצמה והחולשה של הנבחן).

הבעיה הרצינית ביותר של המבחן נעוצה אולי בכך שהניתוח הסטטיסטי (המכונה **ניתוח גורמים**) של מבנה המבחן אינו מאשש את הדרך שבה המבחן בנוי למתן ציונים חלקיים. כלומר, הציונים אינם תואמים את דרכי החשיבה של אנשים, כפי שהן מתגלות בניתוח הסטטיסטי. אפשר, על כן, לקבל לבסוף סדרה של ציונים שאינם אומרים דבר על תהליכי החשיבה של הנבחן. לדוגמה, מה ניתן להסיק על הכשרים המיוחדים של אדם המצטיין בתפיסה מרחבית?

סולמות וכסלר מייצגים חלופה למבחן סטנפורד-בינה, והם סולמות האינטליגנציה הנפוצים ביותר כיום. הם מבוססים על סוג תפיסת האינטליגנציה העומדת בבסיס מבחני סטנפורד-בינה. סולמות וכסלר מחולקים לשלוש רמות: Revised – Wechsler Adult Intelligence Scale (WAIS-R), המהדורה השלישית של ה-Wechsler Intelligence Scale for Children (WISC-III), וה-Wechsler Pre-school and Primary Scale of Intelligence (WPPSI). מבחני וכסלר מניבים שלושה סוגי ציונים: מילולי, ביצועי וכולל. הציון המילולי מבוסס על מבחנים כגון אוצר מילים וכן **דמיון מילולי**, שבהם יש לציין מהות הדמיון בין שני דברים. הציון הביצועי מבוסס על מבחנים כגון **השלמת תמונה**, הדורש זיהוי של חלק חסר בתמונה של חפץ, ו**סידור תמונה**, הדורש סידור מחדש של מערכת מעורבת של תמונות כך שיספרו סיפור לכיד. הציון הכולל משלב את שני הקודמים.

מבחן וכסלר, כמו זה של סטנפורד-בינה, הוא בעייתי במונחים של תוכן. בדמיון המילולי, למשל, הנבחן זוכה לנקודות רבות יותר אם הוא מגדיר את הדמיון בין שני דברים לפי סוגם (כלומר, שניהם שייכים לסוג א), מאשר אם הוא מגדיר אותם לפי תפקודם (לדוגמה, הראשון משתמש בשני) – אף על פי שההכרה בכך שתשובה מן הסוג הראשון "טובה יותר" מן השנייה מייצגת לאו דווקא התפתחות קוגניטיבית, אלא הבנת תהליך המבחן. ניתן היה להבחין בקשר לפי סוג, מבלי להבין שזוהי התשובה המועדפת על הבודן. בדומה לכך, מבחן כמו זה של השלמת תמונה, מניח

ידע של התוכן התמונה; ומבחן של סידור תמונה יהיה בוודאי קל יותר לאלה הרגילים לעקוב אחר רצפים של תמונות המסודרים בסדר זמנים, כגון אלה המצויים בסיפורים מצוירים.

כיצד נבנים מבחני אינטליגנציה

כיצד אלה הכותבים או מפרסמים את מבחני האינטליגנציה בונים את המדדים? שתי גישות יסודיות משמשות לתכנון מבחנים, אך האחת נפוצה בהרבה מן השנייה.

הגישה האמפירית

על פי גישה זאת, נערכת סדרה של תצפיות באנשים המצליחים בהקשר מסוים – כגון בית הספר – ומעוצבות שאלות שיבחינו בין אלה המועדים להצלחה לבין אלה המועדים לכישלון. אפשר גם לבחור בפריטים בהם הסיכוי שילדים מבוגרים יותר ישיבו עליהם נכון גדול יותר מאשר זה של צעירים יותר. כך עשה בינה, כאשר בחר בפריטים שיבחינו בין ילדים מבוגרים לצעירים יותר. גישה זאת נשמעת לכאורה הגיונית – הרי ילדים נעשים חכמים יותר ככל שהם מתבגרים – אך היא מעוררת מספר בעיות רציניות, הגורמות לנו היום לדילמות שאנו ניצבים בפניהן לגבי מבחני אינטליגנציה.

הקושי הראשון הוא בכך שהמבחנים מתבססים על מדגמים הכוללים ילדים בלבד. יש לזכור שהמבחנים שימשו, ועדיין משמשים, להבחנה בין ילדים מחוננים לבלתי מחוננים; כך שהמבחנים מתאימים לילדים הרבה יותר מאשר למבוגרים. כאן נוצרת דילמה. ילדים מבליים זמן רב יותר בפתרון בעיות מילוליות במתמטיקה, בעוד שמבוגרים רבים לא עסקו בבעיות כאלה מזה שנים רבות. כאשר מבוגרים פונים בבקשת קבלה למכללה או ללימודים מתקדמים, הם מצופים לעמוד במבחנים כגון SAT או GRE, הדורשים זכירת עקרונות באלגברה ובגאומטריה שלא עסקו בהם מזה עשרות בשנים. ברור שמבחנים כאלה אינם מודדים אצל

מבוגרים את מה שהם מודדים אצל תלמידי בית ספר או אצל צעירים שזה עתה סיימו אותו. אין פלא שמבוגרים מקבלים לעתים ציונים נמוכים במבחנים כאלה.

המבחנים פותחו, אם כן, על סמך פריטים המתאימים לחומר הלימודים בבתי הספר ולא לזה שהמבוגרים נתקלים בו בחיי היום-יום. נכון היה לומר שהם מודדים אינטליגנציה אקדמית. יתר על כן, חומר הלימוד שהם מבוססים עליו מערבי, וילדים רבים בעולם אינם מעורים בו. גם בארצות הברית קיימים סוגים שונים של בתי ספר, כגון זה של יהודים חרדים הלומדים בישיבות המדגישות פלפול תלמודי ולא את הקנון הספרותי המערבי.

דילמה קרובה לזו היא עצם פעולת הבחינה. ילדים שזה עתה החלו ללמוד בבית הספר אינם מכירים מבחנים. איני מאמין במבחנים לילדים מתחת לכיתה ד (בני תשע או עשר בערך), כי הללו אינם מכירים את אמצעי המבחן. אני זוכר עדיין את מבחן ההישגים הראשון שלי עם גיליון תשובות נפרד, בהיותי בכיתה ג. כאשר הגעתי לשאלה האחרונה בגיליון התשובות של תת-מבחן, גיליתי לתדהמתי שאני נמצא במקום המיועד לתשובה האחת-לפני-אחרונה: במשך המבחן שובש סדר תשובותיי. המורה בוודאי לא תדע זאת, ותחשוב שהידע שלי נמוך ביותר בחומר שאותו תת-מבחן אמור לבחון! במשך שנות הלימודים בבית הספר הילדים מתרגלים עמידה במבחנים, כך שמכשיר המבחן עצמו נעשה מוכר יותר. ייתכן שילדים שלא ביקרו בבתי ספר מערביים לא נתקלו במבחנים כאלה מימיהם. גם מבוגרים, בארצות המערב או בכל מקום אחר, לא ראו מבחנים כאלה זה שנים רבות. חוויית המבחן עצמה עלולה להוות לגביהם מקור של מתח וחרדה. גם כאן, אין לפרש את תוצאות המבחנים באותה דרך כמו לגבי אנשים צעירים.

המחקר מוכיח שנודעת משמעות גם למשתתפים הקשורים לבוחן. ילדים שחורים, למשל, עשויים להיות במצב נחות בנסיבות של מבחן אינטליגנציה טיפוס. גם סוגיות גלויות פחות לעין מעורבות כאן. בקיץ אחד, לפני שנים רבות, פיקחתי על בחינה במסגרת עבודתי ב"חברה הפסיכולוגית" – אחת מן ההוצאות הגדולות של מבחנים, שהייתה אז בניו-יורק. החברה החליטה שכדי להגיע לתקנון אחיד לגבי כל הנבחנים, יש להציג את המבחן באמצעות קלטות של רשמקול. בדרך זאת, ישמעו כל

הנבחנים אותו קול המציג את השאלות באותו קצב ובאותה דרך. הרעיון נראה להלכה כמוצלח, אך למעשה היה זה סיפור שונה לחלוטין. ניתן היה לצפות שהקול בקלטת יהיה זה של קריין מקצועי, המדבר לאט וברור, אולי במבטא של המערב התיכון המובן לכול. אך משום מה, נשיא החברה החליט שהוא עצמו יקרא את המבחנים. הוא בוודאי לא היה קריין מקצועי, אך גרוע מזה – מבטאו הדרומי המתנגן היה אולי מובן בדרום, אך נבחנים ניו יורקים רבים נדרשו למאמץ ניכר כדי להבין את דבריו. גרוע מזה, המחשבה שכולם יהיו מסוגלים לעקוב אחר הקלטת היה מגוחך. התבוננתי במספר רב של נבחנים שאיבדו את דרכם והתקשו לאתר את המקום במבחן בו הם מצויים. הם היו במקום אחד, והקול המכני היה במקום אחר לחלוטין. אך כאשר המעסיקים ששקלו את העסקתם קיבלו את התוצאות, הם ראו רק תדפיס מחשב נקי ומסודר, שעשה רושם שהמבחן התנהל למופת.

חיסרון נוסף של השיטה האמפירית לבניית מבחנים הוא בכך שכותבי המבחנים אינם מחויבים להתעמת עם הסוגיה של מהות האינטליגנציה. הם משתמשים באותם דברים שפועלים – למשל, הגורם המבחין בין ילדים בוגרים לצעירים יותר. ילדים גדולים גבוהים יותר מן הקטנים וגם כבדים יותר; אך עורכי המבחנים אינם שולפים סרגל או מאזניים כשיטה למדידת אינטליגנציה: גובה ומשקל פשוט אינם מקובלים כמדדים של אינטליגנציה. להלכה (שלא כמו למעשה), אין היגיון רב יותר בהכללת אותם פריטים המופיעים במבחנים מאשר בהכללת הגובה והמשקל. ללא תיאוריה של אינטליגנציה אין כל הגדרה של אינטליגנציה, אלא אם כן נקבל את הטיעון החלש ש"אינטליגנציה היא מה שהמבחנים מודדים" – הגדרה מעשית שהוצעה על ידי הפרופסור המנוח באוניברסיטת הרוורד ששמו היה, כיאה, בֹּרִינג [משעמם].

הגישה המבוססת על תיאוריה

הגישה השנייה לבניית מבחנים יוצאת מתיאוריה של אינטליגנציה, ובונה מבחנים על בסיס תיאוריה זאת. היתרון הוא, כמובן, בכך שנקודת המוצא היא תפיסה ברורה המתייחסת למובן של אינטליגנציה. צרכני מבחנים –

פסיכולוגים, בתי ספר או מעסיקים הרוכשים אותם – יכולים להחליט האם התיאוריה מוצאת חן בעיניהם, ובהתאם לכך, האם הם מעוניינים לרכוש את המבחנים. זוהי גישה הוגנת יותר למבחנים. מפרסמי המבחנים מצהירים על משמעות האינטליגנציה לפי תפיסתם, והצרכנים מחליטים בהתאם אם ברצונם לרכוש אותם. זוהי האמת שבפרסום: אומרים לך מה אתה מקבל.

מבחני אינטליגנציה בודדים בלבד בנויים על בסיס תיאוריה של האינטליגנציה. מדוע? כי אימוץ תיאוריה מכריח את המוציאים לאור להתחייב לתפיסה כלשהי של אינטליגנציה. בדרך כלל אין הם מעוניינים לעשות כן, אם משום שאין להם כל תפיסה בנדון, או בשל רצונם להגדיל ככל האפשר את נתח השוק שלהם. הם מניחים – כפי שהניחו יצרני המזון לפני הפעלת החוק של האמת בפרסום – שכלל שייוודע לצרכנים פחות על המוצר שהם מקבלים, כך ייפגעו פחות. כיום, אם אינך אוהב מזון רווי שומן, אתה יכול להחליט שלא לקנות אותו. לפני כן לא ידעת את אחוז השומן, ולכן לא יכולת לדחות פריט מזון על בסיס זה. בכל הנוגע למבחנים, מכיוון שלעתים קרובות אין להם בסיס תאורטי ממשי, אינך יודע בדיוק מה אתה מקבל.

❑ כיצד מעריכים מבחני אינטליגנציה

מובן שצרכני המבחנים אינם אידיוטים מושלמים, והם דורשים מידע על המבחנים המוצעים להם כדי שיוכלו להפעיל שיפוט הגיוני כלשהו בהחלטתם – או לפחות שהדבר ייראה כך. מהם סוגי המידע על המבחנים שהם, ואתה, צריכים לקבל?

תקפות

הסוג הראשון והחשוב של מידע, שיודגש לכל אורך הספר, מכונה **תקפות אמות המידה**. השאלה כאן היא המידה שבה תוצאות המבחן מתייחסות למה שהן אמורות להתייחס אליו. האם המבחן מודד את מה שהוא אמור

למדוד? תקפות של מבחני אינטליגנציה נבחנת אל מול ציוני בית ספר לגבי תלמידים, או אל מול הערכות ביצוע בתכניות הכשרה.

חוברות ההדרכה הטכניות למבחנים מציגות לעתים כמות רבה של נתונים סטטיסטיים הרלוונטיים לתקפות אמות המידה של המבחנים. אך קיימת כאן בעיה: עד כמה, אם בכלל, עלינו לעסוק בציוני בית הספר? האם עלינו לעסוק בהם בהעריכנו באיזו מידה אדם אחד אינטליגנטי יותר מרעהו? אולי עלינו להתייחס לתוצאות של מבחנים מתוקננים של הישגים? בן, למשל, ייסד עיתון חדש בבית ספרו, כותב טורים בעיתון בית הספר, מטיס מטוס ומתרגם לאנגלית תיאורים של בעיות רפואיות לחולים דוברי ספרדית בבית החולים המקומי. אלה הישגים לא רעים לבחור בן שבע עשרה; אך אף לא אחד מהם יופיע בין אמות המידה המשמשות לבדיקת תקפותם של מבחני אינטליגנציה. האם אכן חשובים לנו ציוני בית הספר של אנשים מבוגרים? האם יש להעריך מועמדים למשרות של קברניטי תעשייה לפי ציוניהם בבית הספר? מסופקני.

עצוב לחשוב שאנו מתייחסים לציונים במבחני יכולת כנתונים מוצקים, ולפעילויות מחוץ לתכנית הלימודים כשוליות. בשל הנוחות שאנו חשים בכימות נתונים, אנו יורדים לרמה של ספירת הפעילויות ולא של הערכתן. כתוצאה מכך, תלמיד שמצטרף לשלוש אגודות נחשב למצליח יותר מזה שהוגה, מייסד ומנהל אגודה אחת. המנבא הטוב ביותר של אינטליגנציה מצליחה בעתיד הוא אינטליגנציה מצליחה בעבר. מנבא זה עשוי לשמש בפעילויות בעולם האמיתי המתפתחות לאורך זמן, אך לא בפעילויות בעולם המבחנים המתבצעות במשך שלוש שעות.

אני טוען זה שנים רבות שהמבחנים אינם מודדים את הדברים החשובים לנו באמת לצורך ביצוע – בבית או במקום העבודה. במיוחד טענתי שברמת הלימודים המתקדמים, שבה אנו מכשירים אנשי מקצוע, המבחנים אינם מודדים את המיומנויות החשובות באמת להצלחה בעיסוק כלשהו. לבסוף החלטתי לבחון את אמונתי וחקרתי, יחד עם וונדי ויליאמס (Williams), מהו הדבר שה-Graduate Record Examination מנבא ברמה של הלימודים המתקדמים. בדקנו את כל אלה שהתקבלו ללימודים מתקדמים בפסיכולוגיה במשך שנים עשרה (166 תלמידים), והשווינו את הניבוי של המבחנים עם הציונים בשנה הראשונה והשנייה. יתר על כן, השווינו

אותו להערכות הפרופסורים על התלמידים בתחומים הבאים: (א) כשרים אנליטיים; (ב) כשרים יצירתיים; (ג) כשרים מעשיים; (ד) כשרים מחקריים; (ה) כשרים הוראתיים, (ו) כשרים בכתיבת עבודות מחקר. כמובן, בלימודים מתקדמים משתנים אלה מוערכים כחשובים בהרבה מן הציונים.

מה היו התוצאות? בדקנו גברים ונשים (שהתברר שהישגיהם במבחנים ובלימודים היו ברמה שווה), ומצאנו שהמבחנים ניבאו נכונה את הציונים בשנה הראשונה ללימודים המתקדמים, ולא עוד. רמת הניבוי גם היא לא הייתה גבוהה במיוחד. במבחן המילולי של ה-GRE, ציון המתאם היה 0.18 (על רצף של 0=נמוך, 1=גבוה), שהיה על גבול המובהקות הסטטיסטית, בעוד שבמבחן הכמותי, ציון המתאם היה 0.14 – אפילו לא מובהק סטטיסטית. במבחן הכשרים האנליטיים, היה ציון המתאם 0.17 – שוב על גבול המובהקות הסטטיסטית. שלא במפתיע, המתאם הטוב ביותר היה במבחן ההישגים המתקדם בפסיכולוגיה: 0.37. הישגים במבחן היו המנבא הטוב ביותר להישגים בתכנית הלימודים. המבחן לא ניבא את המדדים החשובים יותר, ואפילו לא את ציוני השנה השנייה. אין תמה אפוא שהמידע על תקפות המתפרסם על ידי המוציאים לאור של המבחנים מדגיש כל כך את ציוני השנה הראשונה!

כאשר בדקנו את התוצאות לפי מגדר (ג'נדר), מצאנו הבדל מעניין. התת-מבחן של כשרים אנליטיים ניבא ברמה נמוכה אך מובהקת אחדים מן המדדים המשמעותיים יותר – אך זאת רק לגבי גברים. המתאם בין ציון הכשרים האנליטיים של ה-GRE לבין הערכת הפרופסורים את כושריהם האנליטיים של הגברים היה 0.31. לנשים, מתאם זה היה 0.05 – כלומר, למעשה לא היה למבחן כל ערך ניבוי לגבי נשים, כך שצירוף גברים ונשים יחד ייתן תמונה מטעה של יכולת המבחן לנבא הצלחה בלימודים מתקדמים. אף על פי שביצועיהם של גברים ונשים בתכנית הלימודים שווים, הרי שהמנבאים הקודמים להצלחה אינם זהים לגבי גברים ונשים. איננו יודעים עדיין מהם ההבדלים במדויק.

ניתן להפיק מן המחקר מספר לקחים. ראשית, המבחנים מנבאים בעיקר ציונים, ולעתים קרובות ציונים בתקופה הקרובה למועד המבחן. אין פסול בניבוי ציונים – נודעת להם חשיבות כי אנו מייחסים להם חשיבות – אך החיים כוללים הרבה מעבר לציונים בלבד. שנית, המבחנים אינם

מנבאים באופן שווה לגבי כל הקבוצות, דבר שאנו נוטים לשכוח כשאנו מעיינים בתוצאות מבחן מבלי להביא בחשבון למי היה מיועד. נשים ניצבות בפני אתגרים אחרים מאשר גברים, ואכן חוקרים מצאו שהערכתן העצמית הנמוכה לגבי ביצוע מטלות מסוימות, גורמת להן להצליח פחות במבחנים. כאשר את מצפה לכישלון, יש להניח שזה מה שתשיגי.

לא היינו מופתעים ביותר מן התוצאות שקיבלנו במחקר על אודות ה-GRE. כמו כן, לא הופתענו מן התגובה לנתונים שהצגנו בפני קבוצות שונות. ספקני המבחנים מהנהנים בהסכמה למשמע תוצאותינו. הגרעין הקשה של המאמינים במבחנים מקבל את נתונינו בספקנות. ייתכן שיטענו שכדי לדעת אל נכון כיצד פועל המבחן, עלינו לבחון גם כיצד היו אלה הנכשלים במבחן מצליחים בלימודים.

מבחינה מסוימת, הם צודקים. היה רצוי לדעת את מידת הצלחתם של אלה שנכשלו במבחן. לדעת, רבים מהם היו מצליחים ביותר. מובן שאיננו יכולים לבדוק זאת, כי אנשים אלה נדחו דווקא בשל תוצאותיהם הנמוכות במבחן. לפיכך, הספקנים לא ישתכנעו עד שנערוך מחקר שאין אפשרות לערכו. ניתן היה אולי לקבל מספר מועמדים מדי שנה מבלי להתחשב בתוצאות המבחן, ולעקוב אחר ביצועיהם. בדרך זאת, לא תהיה הטיה במדגם התלמידים שהתקבלו בשל קבלה על סמך תוצאות המבחן. אכן, הדרך האידאלית לניסוי זה הייתה לקבל מועמדים באופן אקראי. אך מובן שאיש לא יוותר על תוצאות הבחינות למשך שנה או שנתיים, וסגל המורים בוודאי שאינו מעוניין בקבלה אקראית. כך שהשיטה נותרת על כנה, וכל אחד ממשיך באמונתו – שלא תהיה כנראה כל דרך להזימה.

נתונים כגון אלה שקיבלנו במחקר ה-GRE מצביעים על הצורך במידע על תקפות מעבר לסוג המידע הצר המסופק בחוברות ההדרכה הטכניות לבחינות. נדרשות לנו אמות מידה רחבות יותר, ועלינו לדעת את תקפות המבחנים לגבי אוכלוסיות שונות במצבים שונים. אנו עלולים להניח שמבחן שנועד לקבוצה מסוימת ישים גם לקבוצה אחרת; אך האמת היא שאין הדבר כך.

מהימנות

קיים סוג נוסף של מדד שמפרסמי המבחנים והצרכנים מעוניינים בו במיוחד כשהם מעריכים מבחנים. **מהימנות** היא אמת המידה שבה המבחן מודד בעקיבות את מה שהוא אמור למדוד. כלומר, אם תיבחן באותו מבחן פעמיים, האם תקבל אותם ציונים?

מרבית המבחנים הנפוצים מהימנים למדי, שאילולא כן לא היו נפוצים; אך אנו מתכוונים לממוצעים. ציוניהם של אחדים הם בעלי תנודות מעטות, בעוד שציוניהם של אחרים הם בעלי תנודות פראיות. כך ניתן לשפר את התוצאות במבחני SAT או GRE או במבחני כניסה דומים באמצעות ספר או קורס המכין לבחינה.

הימצאותם של ספרים וקורסים כאלה, והעובדה שהם מביאים בממוצע לעלייה בלתי מבוטלת בתוצאות, מעוררת שאלות של שוויון. מה יהיה על אותם אנשים שאינם מודעים לקורסים, או שאין ביכולתם לממן אותם, או שאין להם זמן להשתתף בהם? מובן שמפציי המבחנים יצביעו על כך שהשיפור אינו מכריע, ואכן כך הדבר בממוצע. אך קיימים מקרים בהם השיפור הוא רב, וכאשר אנו עוסקים במיליוני החלטות לגבי מיליוני ציוני מבחנים, הרי שגם הבדל קטן בממוצע ישפיע על תוצאות ההחלטות. לגבי אנשים מסוימים, הספר או הקורס יחוללו שינוי משמעותי, והעובדה שהזדמנויות כאלה מצויות ומנוצלות דיפרנציאלית מביאה לכך שתוצאות המבחנים הן בעלות משמעות שונה לגבי אנשים שונים. אם כן, המהימנות הכוללת של המבחנים עשויה להאפיל על השפעות חשובות לגבי יחידים.

המוציאים לאור מבכרים מבחנים שבהם הפריטים בכל תת-מבחן מודדים פחות או יותר אותה תכונה. הם אפילו מפרסמים נתונים המצביעים על המידה שבה פריטים שונים בתוך תת-מבחן נותנים תוצאה אחידה לגבי התכונה הנמדדת. אף שעקיבות זאת נחשבת ליתרון, יש בצדה מחיר. ככל שפריטים רבים יותר מודדים אותה תכונה, טווח התכונות הנמדדות מצטמצם והולך. העקיבות מושגת לעתים קרובות על חשבון רוחב הבדיקה. אין תמה, על כן, שהמבחנים צרים כל כך. הם מפיקים תוצאות עקיבות המתייחסות לאינטליגנציה בלתי פעילה. כיצד אנו משתמשים במדדים אלה?

השימוש במבחנים

השימוש במבחנים נעשה בדרכים המדגימות את העיקרון של הייזנברג (The Heisenberg Principle): הם משפיעים על היסודות שאותם הם אמורים להעריך. בחברתנו, אלה הנמצאים בראש הסולם (עורכי דין, רופאים, מנהלי חברות, פרופסורים) נוטים להיות בעלי IQ גבוה יותר מאשר אלה בתחתיתו (פועלים יומיים, מנקי בתים, מטאטאי רחובות וכדומה). על כן, IQ מקושר להצלחה מקצועית; אך האם הוא גורם להצלחה זאת?

קל להבחין בסכנות שבטשטוש ההבדל בין סיבתיות לבין מתאם. אנו יודעים, למשל, שרוב תושבי ניגריה הם שחורים, ורוב תושבי נורווגיה הם לבנים; אך אין להסיק מכך שהחיים בניגריה גורמים לתושבים להיות שחורים, והחיים בנורווגיה גורמים להם להיות לבנים. עד כאן חוסר ההבחנה בין סיבתיות לבין מתאם. קיים מתאם, אך כמו תמיד, קיימות שלוש אפשרויות סיבתיות. נניח שקיים מתאם בין שני דברים, כגון אינטליגנציה נמדדת והשמה למשרה. לשם הכללה, נכנה אותם גורם א וגורם ב. ייתכן שגורם א הוא הסיבה לגורם ב; ייתכן שגורם ב הוא הסיבה לגורם א; וייתכן שגורם מדרגה גבוהה יותר הוא הסיבה לשניהם.

אפשר שגם המתאם בין IQ לבין השמה למשרה נגרם על ידי שלושה סוגי מנגנונים. IQ גבוה עשוי בהחלט להיות הסיבה לקבלת משרה טובה יותר, או שמשרה טובה יותר עשויה לגרום ל-IQ גבוה יותר – וגם זה נכון: משרה טובה מאפשרת לאדם לנצל כשרים אינטלקטואליים, ובכך להגביר את האינטליגנציה ולהעלות את ה-IQ. ייתכן גם שהן ה-IQ והן המשרה הטובה תלויות בגורם או בגורמים אחרים. מה עשויים להיות גורמים אלה? כדי להתקבל ללימודי משפטים בארצות הברית, עליך לעמוד במבחן הכניסה המשפטי (LSAT); כדי להתקבל ללימודי מנהל עסקים, עליך לעמוד במבחן הכניסה למנהל עסקים (GMAT). קבלה לבית ספר לרפואה דורשת עמידה במבחן ללימודי רפואה (MCAT). קבלה ללימודים מתקדמים דורשת עמידה במבחן ללימודים מתקדמים (GRE או MAT). לימודים מתקדמים ככלל דורשים עמידה במבחן, ולשם קבלה לתכניות לימודים תחרותיים נדרשים ציונים גבוהים במיוחד. מכללות תחרותיות דורשות מבחן כניסה, וכן בתי ספר פרטיים תחרותיים.

מהו הקשר אם כן בין עובדות אלה לבין המתאם בין IQ לבין השמה למשרה? מבחני האינטליגנציה המקובלים קשורים למבחני הקבלה השונים בדיוק כפי שהם קשורים זה לזה, כלומר – למרות השוני בשמותיהם, כולם מודדים מיומנויות זהות. השימוש בשמות שונים ובתוכן שונה במקצת מועיל אולי לעסקים, אך אינו משנה בהרבה את התוצאות. אלה המצליחים במבחן אחד יצליחו בכל האחרים.

נניח שקיימים אנשים שמסיבה כלשהי אינם מצליחים במבחנים. ייתכן שהם חכמים מבחינה יצירתית אך לא מבחינה אנליטית. ייתכן שהם בעלי חכמה מעשית – יש להם שכל ישר מצוין – אך היא אינה מתורגמת לכישורים אקדמיים מופשטים. ייתכן שהם סובלים מחרדת מבחנים, כפי שסבלתי אני בזמני. ייתכן ששפת אמם היא ספרדית, צרפתית או וייטנמית, ואינם יודעים אנגלית די הצורך. אולי שפת אמם היא אנגלית, אך הוריהם עבדו קשה כל חייהם ולא היה להם זמן לקרוא בפניהם או להעניק להם חשיפה מילולית נאותה. מכל מקום, הם אינם מצליחים במבחנים.

אם אנשים אלה מעוניינים ללמוד בבית ספר יוקרתי לתארים מתקדמים – מן הסוג המספק כרטיס כניסה לעלייה בסולם התעסוקתי – איתרע מזלם, שכן בתי ספר אלה, בכל מקצוע שהוא, דורשים כולם ביסודו של דבר עמידה באותו מבחן. אי לכך, הם יישארו באורח שיטתי מחוץ למעגל הזדמנויות ההשכלה, ובדרך כלל ירדו לבסוף בסולם התעסוקתי. מצד שני, המצליחים במבחנים אינם בהכרח מוכשרים בתחומים רבים, אך יש להם בפירוש כישרון להצליח במבחנים, והצלחה זאת היא יתרון גדול בקבלה ללימודים לתואר מתקדם במכללה לפי בחירתם. הם יתקבלו לפחות לחלק מתכניות אלה. ציוני המבחנים הם אפוא כרטיס הכניסה לרמה המקצועית.

אין זה מפתיע, על כן, שהעומדים בראש הסולם התעסוקתי הם בעלי הציונים הגבוהים במבחנים. הם לא היו יכולים להשתלב במסלולי הגישה לתעסוקה ברמה גבוהה לולא השיגו ציונים גבוהים. למעשה, **אנו עצמנו יוצרים את המתאם בין IQ לבין רמת המשרה.**

חשוב להבין שפני הדברים לא היו כך מאז ומעולם, ואין הם כך ברחבי העולם של היום. בשנות החמישים היו ציוני ה-SAT נמוכים בהרבה מאלה שלאחר עשור. האם האוכלוסייה נעשתה לפתע חכמה בהרבה? יש להניח

שלא. אלא שמכללות רבות שינו את תנאי הקבלה, והחלו להדגיש יותר ציוני מבחנים ופחות את המעמד החברתי. מנקודת ראות זאת, ההישענות על ציוני המבחנים העניקה לרבים הזדמנות טובה יותר להתקבל על סמך כישוריהם האינטלקטואליים. בארצות אחרות עדיין משפיעים גורמים אחרים, כולל המעמד החברתי, על הנגישות ללימודים גבוהים.

הבה נבחן אנלוגיה מסוימת: נניח שאנו כחברה מחליטים שציוני המבחנים אינם מעניינים אותנו, כי תוצאות המבחנים משתנות כל כך; ניתן להיבחן ב-SAT ולהשיג יום אחד ציון 500, ולמחרת – 570. על כן אנו מחליטים על אמת מידה עקיבה יותר – הגובה, למשל. מעתה, קבלה למכללה וללימודים לתואר מתקדם תיקבע לפי הגובה. עליך להיות בגובה של 2 מטר כדי להתקבל להרוורד. כדי להתקבל ליל, מספיקים 1.98 מטר (בתנאי שאתה יפה תואר). כך נמשיך, עד שנגיע לאוניברסיטה המצויה בתחתית המדרגה, המסתפקת לצורך קבלה בגובה של מטר ורבע בלבד. מובן שכדי להתקבל לבית ספר תחרותי לרפואה או למשפטים יידרש גובה של 2.10 מטר. דבר זה אינו מגוחך לחלוטין, כפי שנראה לנו ממבט ראשון. עובדה היא שאלה העומדים בראש סולם התעסוקה נוטים להיות גבוהים יותר מאלה שבתחתיתו. ובכן, גובה נחשב – בין אם אנו מודים בכך בין אם לאו.

עשרים וחמש שנה לאחר תחילת השימוש בגובה כמשתנה הקובע החלטות קבלה, אנו מחליטים להשוות את ממוצע ה-IQ של אנשים במקצועות שונים – עורכי דין, רופאים ומנכ"לי חברות מצד אחד, ופועלים יומיים, עובדי ניקיון, ועובדים בקו הייצור מצד שני. נגלה שככל שהאדם נמצא גבוה יותר בסולם המקצועי, קומתו גבוהה יותר. האם הוכחנו בכך שגובה מהווה יתרון בקבלת עבודה טובה או משרה רמת דרג? לא. מה שהוכחנו הוא שהשתמשנו בגובה כבסיס להחלטה באילו משרות אנשים יורשו – או לא יורשו – להשתבץ.

אין כוונתי לומר ש-IQ אינו קשור להצלחה בעבודה. הוא אכן קשור, אם גם לא בקשר הדוק. כוונתי היא שאל לנו להסיק, כפי שעושים זאת הרנסטיין ומורי (Herrnstein and Murray) בספרם השנוי במחלוקת **עקומת הפעמון**, שישנה יד מכוונת כלשהי – כוח טבעי כלשהו – האחראית לכך שהשמנת צפה כלפי מעלה והפסולת שוקעת לקרקעית. ההפך הוא הנכון:

עלינו להכיר בכך שנקבל כתברה את מה שאנו יוצרים. אין זה הטבע המחליט את מי להעריך; החברה היא המחליטה זאת.

ה-IQ אמנם נחשב, אך לא במידה רבה. מדדים המבוססים על IQ מסבירים פחות מ-10 אחוז מן השונות בין אלה המצליחים יותר לבין אלה המצליחים פחות לפי קני המידה החברתיים. פירוש הדבר שה-IQ מותר ללא הסבר יותר מ-90 אחוז מן השונות בין בני אדם. מכאן שלא ניתן לטעון על סמך מדדים אלה שה-IQ הוא הקובע. אולם, האם יש בסיס לטענה בדבר הבדלים ב-IQ בין קבוצות?

הבדלים קבוצתיים ב-IQ ובאינטליגנציה

דבר אחד ברור: בארצות הברית, קבוצות גזעיות ואתניות הן בעלות ממוצע IQ שונה. עובדת השוני כשלעצמה אינה מהווה הוכחה לכך שהמבחנים מוטים. ההטיה אינה פונקציה של הבדלי ציונים בין קבוצות נבחנים, אלא של ניבוי. הבדלים בין קבוצות היו מהווים הוכחה להטיה רק במקרה שמה שהמבחן אמור לנבא אינו מציג הבדלים דומים.

הספרות על הטיה במבחנים נסקרה בהרחבה עד לזרא. המסקנה הכללית מספרות זאת היא שהמבחנים אינם מוטים, לפחות לא במובן במקובל. עלינו להרהר במשמעותו של ממצא זה.

פירושו של הממצא הוא שאם קבוצה א מצליחה פחות מקבוצה ב במבחן אינטליגנציה, הרי שהיא תצליח פחות – באותה מידה לערך – גם בביצועים שאותם המבחן אמור לנבא, אם אמנם המבחן מנבא ביצועים אלה. על כן, אם קבוצה א מצליחה פחות מקבוצה ב במבחן האינטליגנציה המקובל, הרי שהיא תצליח פחות גם בהישגים לימודיים בבית הספר. שחורים בארצות הברית, למשל, מקבלים ציונים שהם בסטיית תקן אחת (15 נקודות IQ) בממוצע פחות מן הלבנים. הישגיהם בבית הספר גם הם נמוכים יותר. הרי לכם הוכחה – על פי ארתור גינסן, ריצ'רד הרנשטיין וצ'רלס מורי ואחרים מן הימין האינטלקטואלי – שהמבחנים אינם מוטים. אמריקנים ממוצא יפני או אסיאתי אחר נוטים להצליח מעט יותר

מן הלבנים במבחנים רבים, וגם בלימודים בבית הספר. הרי לכם – אין הטיה! אך המתינו נא רגע.

האם ההבדלים בין הקבוצות הגזעיות תורשתיים? הרנשטיין ומורי משיבים בחיוב על שאלה זו,⁵ אך הנתונים אינם מאשרים את מסקנתם. הם ואחרים מצביעים על התורשתיות של האינטליגנציה כמרמזת על התורשתיות של הבדלים בין קבוצות. למעשה, הנתונים הקיימים הם מה שקרוי הערכות של תורשתיות תוך-קבוצתית. הם מורים לנו על מקורות ההעברה של האינטליגנציה בתוך קבוצות אך לא ביניהן. המחקר על תאומים לבנים מלמד על מקורות השוני בין פרטים שהם תאומים לבנים, אך לא על מקורות השוני בין קבוצות של פרטים שחלקם לבנים וחלקם שחורים.

ההבדל אינו עניין של דקדקנות סטטיסטית בלבד. דוגמה נפוצה שהרנשטיין ומורי עצמם משתמשים בה: נניח שיש לנו חופן של זרעי תירס המראים את השונות הטיפוסית לגרעיני תירס. נזרע את מחציתם בשדה פורה ואת המחצית השנייה בשדה שחון במדבר. אם כי תכונות התירס הן תורשתיות, ההבדלים בהתפתחות בין שתי הקבוצות תלויים לחלוטין בסביבה. כיצד ניתן ליישם היגיון זה להבדלים בין שחורים ולבנים?

גם אם האינטליגנציה עוברת בתורשה במידה מתונה, הרי שהתורשתיות כפי שהיא נקבעת בתוך קבוצות אינה מלמדת דבר על הסיבות לשוני בין קבוצות. זוהי משמעות הדוגמה של התירס. יתר על כן: בהשוואה בין קבוצות, עלינו להגדיר היטב את הקבוצות שאנו משווים. יהיה זה אוילי לדבר על גזעים טהורים. אמריקנים ממוצא אפריקני מייצגים בעיקרו של דבר תערובת בין פרטים שחורים ממוצא אפריקני לבין פרטים לבנים ממוצא אירופי ואחר. קבוצות הגזע המשמשות למחקר פסיכולוגי מורכבות מקבוצות חברתיות, ולא ביולוגיות. כלומר, בני אדם מוגדרים כשייכים לגזע מסוים על פי דבריהם, ולא על פי מוצאם מלידה.

העדויות מרמזות על כך שהבדלים בין שחורים ללבנים הם בעיקר סביבתיים, ולא גנטיים. מאות אחדות של ילדים גרמנים שאבותיהם היו חיילים אמריקנים שחורים במלחמת העולם השנייה, היו בעלי IQ ממוצע דומה עד כדי מחצית הנקודה לזה של ילדים שאבותיהם היו חיילים אמריקנים לבנים. זאת ועוד, ילדים מנישואים מעורבים של שחורים

ולבנים הם בעלי IQ גבוה יותר אם אמותיהם לבנות – דבר המעיד על השפעת הסוציאליזציה ולא הגנטיקה.

זאת ועוד: ריצ'רד ניסבט (Nisbett), פסיכולוג באוניברסיטת מישיגן, סקר את הספרות ומצא שבעה מחקרים המשווים מוצא גנטי עם מוצא סביבתי ביחס להבדלים ב-IQ בין שחורים ללבנים. מחקרים אלה, שלא כמו מחקרי התאומים, ביקשו למצוא את המקור הישיר של ההבדלים בין שחורים ללבנים. שישה מהם לא הצליחו למצוא כל הוכחה להשפעות גנטיות. מחקרים של סקאר וויינברג (Scarr and Weinberg, 1976) הניב תוצאות מפורקפות. אין זה ברור כיצד יש לפרש את הממצאים, אם כי החוקרים עצמם מפרשים אותם כלא תומכים בהסבר הגנטי. מעניין ומדכא לציין שהרנשטיין ומורי דנים בהרחבה במחקר דו-משמעי זה בלבד, ומובן שפירושה את התוצאות הפוך מזה של המחברים עצמם. כאן, כמו במקומות אחרים, הם מתעלמים מעדויות שאינן תומכות בהשערותיהם, ומפרשים עדויות דו-משמעיות בצורה שתתמוך בדעותיהם הפוליטיות-החברתיות.

האם ה-IQ הוא העומד מאחורי הבדלים קבוצתיים בסוגים שונים של הצלחה חברתית? הרנשטיין ומורי חושבים שכך הדבר, וספרם מבקש להבהיר עמדה זאת. אין הם לבדם: גם ג'נסן ואחרים מאמינים כמותם. אך למרבית הפלא, נתונייהם של הרנשטיין ומורי עצמם אינם תומכים בטענתם. בסדרה אחת של ניתוחים, הם השוו שחורים ולבנים בעלי ממוצע IQ זהה. הם מצאו שהסבירות של השחורים לחיות בעוני כפולה מזו של הלבנים; שהסבירות שייולדו מחוץ למסגרת הנישואים גבוהה פי חמישה; שהסבירות שיהיו נתמכי סעד גבוהה פי שלושה; שהסבירות שיחיו בעוני בשלוש השנים הראשונות לחייהם גבוהה יותר מפי שניים, ושהסבירות שייולדו במשקל נמוך גבוהה פי שניים. ממצאים אלה אינם תומכים בהנחה שה-IQ הוא הגורם להבדלים בהיבטים שונים אלה של הסתגלות, מכיוון שה-IQ של שתי הקבוצות היה זהה.

מאז תחילת המאה היו כאלה שהתריעו על הבדלי קבוצות ב-IQ, והזהירו מפני הסכנה הטמונה בהבדלים אלה – לא בשל עצם קיומם אלא בשל כך שקצב הילודה השונה בין הקבוצות יביא לירידה ברמת האינטליגנציה הלאומית. אם אכן אלה המדורגים נמוך בסולם החברתי-הכלכלי ובסולם ה-IQ הם בעלי קצב ילודה מהיר יותר – הרי

שה-IQ ירד בין אם האינטליגנציה תורשתית או סביבתית, שהרי הגנים הרעים והסביבה הרעה שיסופקו על ידי ההורים הרעים בעלי ה-IQ הנמוך יביאו לסחף כלפי מטה.

השערה זאת מקובלת על המתריעים מפני התופעה, למרות העובדה שבמשך שלושים השנה האחרונות לערך עלה ה-IQ בצורה מרשימה, בסטיית תקן שלמה (כ-15 נקודות IQ). המתריעים מתקשים מאוד בקשר לעובדה זאת המכונה "אפקט פלין" ("Flynn effect") – על שם ממציאו ג'יימס פלין⁶ ואינם מסוגלים להסבירה. אפקט פלין אינו מוגבל לארצות הברית בלבד, אלא נפוץ במדינות רבות בעולם. הישגים שנחשבים היום לרמת IQ של 85 נחשבו לפני שני דורות בלבד לרמה ממוצעת – IQ 100.

הוצעו השערות רבות בנוגע למקורו של אפקט פלין, אך לא התקבלו כל מסקנות. השכלה טובה יותר, תזונה טובה יותר, חינוך ממושך יותר – כולם בגדר האפשר. הדבר היחיד הברור הוא שלמרות התחזית על סחף כלפי מטה, הכשרים הממוצעים ממשיכים לעלות.

עם זאת, בתחום אחר קיים סחף כלפי מטה. במקביל לעלייה המתמדת בציוני ה-IQ, הציונים במבחני קבלה שונים למכללות נמצאים בירידה. מובן שמבחנים אלה מודדים רק חלק מן הכשרים. עם זאת, הסימנים מצביעים על ירידה במיומנויות אקדמיות. פרופסורים רבים מציינים, כמוני, סחף כלפי מטה במיומנויות המילוליות של תלמידים. סחף זה היה משמעותי עד כדי כך שאחרי שנים רבות שבהן נשמרו הנורמות במבחני ה-SAT, "שירות המבחנים החינוכיים" (Educational Testing Service) קבע נורמות חדשות, והעמיד את הציון הממוצע על 500.

הציונים הנמוכים אינם נובעים, כפי שניתן להניח, מירידת הכישורים של הנבחנים שבתחתית הסולם. אף על פי שמתקבלים כיום ציונים נמוכים רבים יותר, מקור המגמה הוא דווקא בהבדלים בין אוכלוסיות הילדים הנבחנות. תלמידים רבים יותר נבחנים כיום מאשר בשנות החמישים, השישים או אפילו השבעים, כאשר הציונים היו גבוהים יותר. באותן שנים נבחנו התלמידים הטובים בלבד; היום נבחנים תלמידים מכל הסוגים. הבעיה האמיתית היא הירידה בציונים הגבוהים. עד לקביעת הנורמות החדשות, הייתה ראויה לציון הירידה בציונים בקצה העליון של הסולם – ברמת ה-700 וה-600. מדוע ירדו ציונים אלה? נראה לי שיש לכך שלוש סיבות.

הסיבה הראשונה היא העיסוק הלאומי הגובר בטעונו הטיפוח. אלה הנמצאים בקצה התחתון של סולם הכשרים ראויים אכן לשירותים מיוחדים, אך כך גם אלה שבקצה העליון. הקדימויות הלאומיות שלנו מתגלות בעובדה ש-99.9 אחוז מן התקציב לחינוך המיוחד מופנה לקצה התחתון של הסולם. זהו שיקול מטופש. הילדים המוחננים – כוונתי לכל הכישורונות המרכיבים את האינטליגנציה, לא רק ל-IQ – הם קרוב לוודאי המשאב הלאומי היקר ביותר שלנו, תקוותנו העיקרית בעולם של תחרות גוברת. ברמת בית הספר היסודי והתיכון, אנו מקדישים לתלמידים המוכשרים תשומת לב פחותה בהרבה מאשר למתקשים. ההנחה היא שילדים מוחננים יסתדרו בעצמם ואינם זקוקים לשירותים. זוהי טעות. הם זקוקים לשירותים לא פחות מאלה המצויים בקצה התחתון. סוג ההוראה שהם זוכים לו משעמם אותם לעתים קרובות, וגורם להם לשנוא את בית הספר. אחד מילדיי, בהחלט לא גאון מתמטי, בילה שני שלישים משנת הלימודים בחזרה על המתמטיקה ששלט בה זה מכבר. זהו בזבז זמן וחרפה. ילדים אלה לא ילמדו לעולם כיצד לנצל את כישורונותיהם כראוי.

שנית, אנו מורידים את רמת בתי הספר, דבר שהוכח על ידי סלי רייס (Reis) מאוניברסיטת קונטיקט, במחקר שנתמך על ידי משרד החינוך של ארצות הברית. ספרי הלימוד לכל כיתה נמוכים בשלוש רמות לערך מרמתם של הספרים שרווחו לפני שלושים שנה כאשר למדתי בבית הספר התיכון. כך גם ספרי הלימוד בבתי הספר היסודיים. כמחבר של ספרי לימוד, אני מוצא שלא מופעל עלי כל לחץ להעלות את רמת הספרים; הלחץ הוא תמיד להוריד את הרמה.

קל להאשים את הוצאות הספרים. מדוע הם עושים זאת לילדינו? הבעיה אינה בהן, אלא בבתי הספר. ההוצאות מדפיסות מה שבתי הספר רוכשים. אם בתי הספר ידרשו ספרי לימוד ברמה גבוהה, הן יפיקו אותם; אם הם ידרשו תמונות צבעוניות נאות – הם יקבלו אותן. מטרת ההוצאות להרוויח, והן מספקות את דרישות הצרכנים.

אנו בארצות הברית מדברים על סטנדרטים גבוהים, והממשלה והקרנות מפיקות דוחות לאין ספור הקוראים להשגתם. אלה הם דיבורים סרק, שאין לצדם מעשים. כאשר מופיעים ספרי לימוד ברמה גבוהה, אין להם שוק. השוק דורש רמה נמוכה, ותעשיית הספרים יודעת

זאת היטב. בתי הספר, ואפילו המכללות, רוכשים ספרים קלים להבנה, כדי שלא להביך את תלמידיהם; אך התלמידים מחמיצים בכך את האתגר. הצביעות אינה מוגבלת לחינוך. ביקרתי בשני תריסר מדינות, ולמיטב ניסיוני אין מדינה שבה חיים אנשים בעלי משקל עודף רבים כל כך כמו בארצות הברית. כיצד אנו מתמודדים עם כך? אנו מעלים את רמת המשקל המקובלת לכל גיל. בינתיים, אנשים מתים יותר מאשר אי פעם ממחלות לב. מה עשו יצרני הבגדים כאשר גילו שהמידות הקטנות אינן נמכרות, ושנשים רבות יותר קונות בגדים במידות גדולות, דבר הגורם להן לדכדוך? הם שינו את סימון המידות, כך שנשים קונות מידות קטנות יותר לכאורה, ומרגישות טוב יותר. ניתן לשנות את התוויות, אך אין זה משנה את מהות הדברים, או את הבעיה. ניתן לקרוא לספר לימוד "רמת כיתה ו", אך רמת חומר הקריאה שבו נמצאת בירידה בכל הכיתות.

בספר **פער הלמידה**, הצביעו הרולד סטיבנסון וגיימס שטיגלר (Stevenson, H. & Stigler, J., *The Learning Gap*) על פרדוקס. בהשוותם הורים יפנים ואמריקנים, מצאו שההורים האמריקנים מרוצים יותר מהישגי ילדיהם מאשר ההורים היפנים, למרות ההישגים הנמוכים יותר באורח ניכר של ילדיהם. הורים אמריקנים מאמינים שהכל מתנהל כסדרו – הכל בסדר בבתי הספר, אך קיימת קונויה להציג את בתי הספר באור שלילי. גם היקף המותניים שלנו בסדר. יכולתנו הלאומית לאשליה עצמית מדהימה. בארצות אחרות לא מטפחים אשליות כאלה. עלינו להתחיל לאתגר את שכלם של ילדינו, ולא לשנות את הנורמות של המבחנים ולהשלות את עצמנו שילדינו קוראים. ברמת המתאימה כאשר הם קוראים חומר במשקל נוצה. מחיר האשליה העצמית הוא ירידה ברמה הלאומית של התחרותיות. זהו מחיר יקר, והוא עולה עם הזמן. עלינו להעלות את רמת האינטליגנציה, לא להוריד אותה.

יכולת שינוי של אינטליגנציה

אינטליגנציה, ואפילו IQ, ניתנים לשינוי. בעבר האמינו שהאינטליגנציה קבועה – אנו תקועים לנצח עם רמת האינטליגנציה שנולדנו איתה. כיום,

רובם המכריע של החוקרים בתחום האינטליגנציה מאמינים שהיא ניתנת לעיצוב, ואף ניתנת לשיפור על ידי התערבות בדרכים שונות.⁷

תכנית Head Start פותחה בשנת 1960 כדי לספק לגיל הרך יתרון בכשרים אינטלקטואליים ובהישגים עם כניסתם לבית הספר. מחקרי מעקב ארוכי טווח מורים על כך שבשנות ההתבגרות, רמת הילדים שהשתתפו בתכנית הייתה גבוהה בשנת לימודים שלמה מרמתה של קבוצת ביקורת של ילדים שלא השתתפו בתכנית.⁸ הילדים שהשתתפו בתכנית קיבלו ציונים גבוהים יותר במגוון מבחנים לבדיקת הישגים לימודיים, היו זקוקים לפחות סיוע מתקן וסבלו פחות מבעיות התנהגות. כל אלה אינם מדדים טהורים של אינטליגנציה, אך הם מגלים מתאם חיובי עם מבחני אינטליגנציה מקובלים. מספר תכניות חדשות הצליחו בסביבות שמחוץ לביתו של הילד. הפרויקט של ריימי וקמפל (Ramey and Campbell, Abecedarian Project, 1984, 1992) שהופעל בצפון קרוליינה הוא דוגמה ראויה לציון. הוא הצליח להעלות את הכשרים האינטלקטואליים של תלמידים במספר מקרים שונים.

עבודתם של רוברט ברדלי ובטי קלדוול (Bradley and Caldwell) הצביעה על החשיבות שנודעת לסביבה הביתית בהתפתחות האינטליגנציה של ילדים צעירים. חוקרים אלה מצאו שמספר גורמים בסביבת הבית המוקדמת (טרום בית הספר) עשויים להיות קשורה לציוני IQ גבוהים: תגובתיות רגשית ומילולית של המטפל העיקרי ומידת עיסוקו בילד; הימנעות מהגבלות שרירותיות ומעונשים; ארגון הסביבה הפיזית וסדר היום של הפעילויות; הספקת חומרי משחק נאותים, והזדמנויות לשינוי וגירויים מתמידים. המחקרים העלו שגורמים אלה ניבאו את ציוני ה-IQ באופן אפקטיבי יותר מאשר המשתנים "סטטוס חברתי-כלכלי" או "מבנה המשפחה" – כגון מספר הילדים.

לסביבה הביתית נודעת השפעה רבה במקרים של חסך קיצוני. הפסיכולוג האמריקני ויין דניס (Dennis), שעבד באירן, הראה שילדים מבתי יתומים מסוימים שלא אומצו עד לגיל שנתיים הפכו ברובם למפגרים שכלית. אלה שאומצו לא היו מפגרים. לא הייתה כל התערבות מיוחדת; ההבדל היה בכמות הגירויים האינטלקטואליים והטיפול של הסביבות השונות. נמצאו כאן הבדלים ב-IQ עד כדי 50 נקודות. מסתבר כי כאשר קיים פער רחב בטיב הסביבה, היא הופכת לרכיב רב-עוצמה.

למעשה, אפילו סביבה הגונה הופכת לאמצעי בעל עוצמה להעלאת ה-IQ, כמו גם היבטים אחרים של האינטליגנציה.

תכנית מרשימה ביותר להדרכה במיומנות אינטלקטואלית היא "פרויקט אודיסאה" (Project Odyssey) בוונצואלה. במשך בתקופה התקיים בה משרד ממשלתי ייחודי שהוקדש כולו לפיתוח האינטליגנציה – יוזמה חסרת תקדים בהיסטוריה האנושית.⁹ התכנית עסקה במגוון רחב של מיומנויות אנליטיות ויצירתיות, ולוותה בהערכה קפדנית. הערכה זאת ראויה לשמש מודל להערכת תכניות כאלה, בהשתמשה במגוון רחב של מדדים קוגניטיביים לקביעת ההצלחה של התכנית. התוצאות פורסמו באחד מכתבי העת הפסיכולוגיים היוקרתיים ביותר.

הזכרתי תכנית זאת במיוחד בשל מחבריה של ההערכה הקפדנית – ריצ'רד הרנשטיין ועמיתיו (Hernstein et al.). המחקר נעשה כתוצאה מן הפרויקט שמטרתו הייתה לפתח בתלמידי בתי הספר בוונצואלה מיומנויות אינטלקטואליות. הרנשטיין היה זה שלאחר מכן, בספרו **עקומת הפעמון**, טען שמאמצים להגברת אינטליגנציה נחלו כישלון חרוץ – דוגמה להצהרות מנופחות בלתי נכונות.

מצד אחד, מתמיה לחשוב שמי שהיה שותף ליצירת תכנית מצליחה והמחבר הראשי של הערכתה יכול לאחר מכן להסיק שתכנית זאת נדונה מראש לכישלון. מצד שני, אין זה מפתיע כלל. תחומי מדע שונים הם בעלי הקשרים פוליטיים חזקים עד כדי כך שוויכוחים בעלי יומרה מדעית הם לעתים קרובות פוליטיים למעשה. חקר האינטליגנציה הוא בבירור תחום כזה, במיוחד בסוגיות של יכולת שינוי והבדלים קבוצתיים. הטענות הפוליטיים באים מן הימין ומן השמאל, והסכנה שבפוליטיזציה היא בכך שהסוגיות המדעיות נעלמות בתוכה.

עלינו להבחין בין הוראה אמיתית של מיומנויות אינטלקטואליות לבין הכנה שגרתית לבחינות. קיימים ספרים וקורסים שבעזרתם ניתן לשפר את ציוני המבחנים, כי הם מספקים אימון בסוגי הפריטים המצויים במבחנים; אך הם אינם תורמים בהכרח לשיפור כולל באינטליגנציה. איני מתנגד לספרים או לקורסים אלה. ילדי מתקרבים לגיל שבו הם ייבחנו ב-SAT, ולפיכך קנתי להם ספר כזה, ואשמח אם ישתתפו בקורס מתאים. אין בכוונתי להעלות את רמת האינטליגנציה

שלהם, אלא לשפר את ציוני המבחן שלהם כך שיתקבלו למכללה שהם יבחרו בה. קורסים אלה מביאים לעלייה קטנה עד בינונית בציוני המבחנים המתוקנים.¹⁰ בעלי אינטליגנציה מצליחה אינם מקבלים בהכרח ציונים מבריקים במבחנים אלה, אך הם מודעים לכך שעליהם להשיג ציונים טובים כדי להמשיך בדרכם במערכת החינוך ועל כן הם משתדלים להשיג ציונים גבוהים ככל האפשר – לא בשל מה שהמבחנים מודדים אלא בשל החשיבות (השגויה) שהחברה מייחסת לציונים אלה.

במקביל, קיימת גם שאלת השוויון בהקשר לספרים וקורסים אלה. יש כאלה שאינם יכולים להקדיש את הזמן והכסף הנדרשים להם, והם נמצאים בעמדה נחותה, בממוצע, ביחס לאחרים. זאת ועוד, בעלי הזמן והממון אינם מפולגים אקראית על פני האוכלוסייה, אלא יש להניח שהם בני המעמד החברתי-הכלכלי הבינוני עד הבינוני-הגבוה. גם כאן, המבחנים משרתים בממוצע את אלה שמצבם שפיר יותר מלכתחילה.

הממצא הרווח – ביחס למבחנים ולמדדים אחרים של הצלחה בחברה – מכונה לפעמים "אפקט מתיו" (Matthew Effect). בברית החדשה, הצביע מתי על כך שהעשירים נוטים להפוך לעשירים יותר והעניים נוטים להפוך לעניים יותר; ואין האמור בכסף בלבד. "אפקט מתיו" נוטה להביא לקיטוב במשאבים החברה – דבר שאנו חוזים בו כיום בארצות הברית ובארצות אחרות – וגם להגברת אי-ההבנות בין אלה שיש להם לבין אלה שאין להם.

בתי הספר, במיוחד התיכוניים, הופכים עם הזמן לכיתות מהוללות של הכנה לבחינות. ביקרתי ב"יום הורים" בשיעור לאנגלית של בתי, והתבוננתי כיצד התלמידים משננים ולאחר מכן מדקלמים הגדרות של מילים קשות מאוצר המילים הנדרש. השיעור שימש מסווה להכנה לקראת ה-SAT. קשה להאשים את בתי הספר בלבד: ההורים, וכן רשויות החינוך המחוזיות, לוחצים עליהם לשפר את הישגי תלמידיהם במבחנים. התלמידים שוכחים לרוב את המילים מיד לאחר שהם משננים אותן, כי הן אינן מעוגנות בתודעתם, והם אינם לומדים כיצד להשתמש בהן וכיצד לקשר ביניהן לבין ידע קודם.

ערכנו מחקר שנראה כעין גישה חלופית לשיפור המיומנות של רכישת אוצר מילים.¹¹ במקום לדרוש מן הילדים ללמוד מילים בעל פה,

הדרכנו את התלמידים בקבוצות הניסוי כיצד להבין את משמעות המילים מתוך התוכן. כאשר נתקלו במילה חדשה בהקשר הטבעי, הם הונחו כיצד להשתמש ברמזים מתוך התוכן כדי להבין את משמעותה. במחקר היו שתי קבוצות ביקורת: האחת קיבלה את סוג האימון המקובל בספרי הכנה למבחנים, אך ללא הדרכה כיצד להגיע למובני המילים באמצעות התוכן; השנייה לא זכתה לטיפול כלשהו. כאשר השווינו את הצלחת הקבוצות, מצאנו שקבוצות הניסוי השתפרו באופן משמעותי יותר מאשר שתי קבוצות הביקורת, במדדים של כושר ההבנה של מילים מתוך הקשרן. שתי קבוצות הביקורת לא נבדלו זו מזו. כלומר, האימון לכשעצמו אינו אפקטיבי במיוחד.

נדרשת הדרכה כדי לדעת כיצד להבין פירוש מילים. זמן רב לאחר שיסימו את בית הספר, התלמידים יהיו בעלי מיומנויות שניתן ליישמן ללימוד מילים חדשות, במקום לשנן רשימות מילים שיישכחו תוך זמן קצר. אם ברצוננו לפתח את האינטליגנציה, עלינו ללמד את תלמידינו לחשוב בהקשר תוכני, ולא רק לשנן. העדויות מורות אכן על כך שניתן ללמד חשיבה טובה יותר.¹² אם נמשיך להסתמך על ה-IQ ועל מגוון המבחנים האחרים כמדדים היחידים של אינטליגנציה, נשלול מרבית מן התלמידים את ההזדמנויות וסוג ההוראה החיוניים לשיפור האינטליגנציה.