מרכז ארצי לבחינות ולהערכה (ע"ר)
NATIONAL INSTITUTE FOR TESTING & EVALUATION
المركز القطري للامتحانات والتّقييم
מיסודן של האוניברסיטאות בישראל

היעילות והתקפות של בחינה ממוחשבת אדפטיבית (מפע"ם) כהתאמה לנבחנים עם הפרעות קשב וריכוז (ADHD)

דו"ח
RR
16-01

ספטמבר 2016

נועה סקה | מיכל באומר | טלי רוזנברג | תמר מלינוביץ

דוח מַאלו 01–16–RR

ISBN:978-965-502-201-8

© כל הזכויות שמורות
מרכז ארצי לבחינות ולהערכה

© All rights reserved
NITE

היעילות והתקפות של בחינה ממוחשבת אדפטיבית (מפע"ם) כהתאמה לנבחנים עם הפרעות קשב וריכוז (ADHD)

נועה סקה
מיכל באומר
טלי רוזנברג
תמר מלינוביץ

ספטמבר 2016

הכותבות מבקשות להודות למרים איתן על תרומתה לתכנון הדו״ח, לעיבודו ולכתיבתו, על הרעיונות הרבים, על המשוב ועל הביקורת; לד״ר צור קרליץ ולד״ר תמי קנת-כהן על הסיוע בבניית מערך המחקר; ליונתן סער על הייעוץ והמשוב לאורך תהליך הכתיבה; ליהודית מירוז ולדורית דהן על הסיוע הטכני; ולכל חברות צוות היחידה לבחינות מותאמות על התגייסותן לסיוע בכל שלבי המחקר – מהעבודה הלוגיסטית הרבה שהייתה כרוכה במחקר ועד למשוב ולסיוע בכתיבת הדו״ח.

תוכן עניינים

תמצית.....4
מבוא.....5
מטרות המחקר.....13
שיטת המחקר.....15
ממצאים.....21
תוצאות משוב מפע"ם.....29
דיון.....41
מקורות.....47

לוחות

לוח 1 – מערך קבוצות המחקר והביקורת בניסוי.....18
לוח 2 – ציוני הנבדקים משתי קבוצות הניסוי בבחינה הניסיונית ובבחינה האמתית.....21
לוח 3 – מתאמים בין הציונים של נבדקי הניסוי בבחינה הניסיונית ובבחינה האמתית.....22
לוח 4 – ציונים מתוקנים בבחינה הניסיונית.....22
לוח 5 – לוח ניתוח שונות לבדיקת השפעת סוג הבחינה והקבוצה הציונים בתחומים השונים ועל הציון הכולל.....23
לוח 6 – טעות המדידה באומדן היכולת (PVAR*) בבחינת מפע"ם בשתי הקבוצות.....25
לוח 7 – מספר המתכנסים ואחוז ההתכנסות בבחינת מפע"ם בשתי הקבוצות*.....26
לוח 8 – מספר הפריטים שקיבלו הנבחנים בבחינת מפע"ם בשתי הקבוצות.....26
לוח 9 – נתונים לגבי תדירות השימוש במחשב ומידת בהירות ההסברים בבחינה הממוחשבת לפי קבוצה.....29
לוח 10 – מידת המאמץ שהשקיעו הנבחנים בבחינה הניסויית הממוחשבת לפי קבוצה.....30
לוח 11 – מידת הקושי הנתפסת של פריטי הבחינה לפי תחומים ולפי קבוצות.....30
לוח 12 – ממוצעים וסטיות תקן של מידת השיפור, לפי העדפת הבחינה.....35

איורים

איור 1 – ציוני קבוצות הניסוי והמחקר בבחינה הניסויית (מפע"ם מול נייר ועיפרון).....24
איור 2 – ציוני קבוצות הניסוי והמחקר בבחינה האמיתית על פי סיווגן בניסוי.....24
איור 3 – הבדלים בין נבחנים שנקבע שציונם ישתפר במפע"ם לבין נבחנים שנקבע שציונם לא ישתפר.....27
איור 4 – התפלגות ציוני הבחינה לפי חלוקת שיפוטי הקוראות.....28
איור 5 – התפלגות שביעות הרצון מן הצפייה הממושכת במסך בבחינת מפע"ם.....31
איור 6 – התפלגות שביעות הרצון מהצגת כל שאלה במסך נפרד בבחינת מפע"ם.....32
איור 7 – התפלגות שביעות הרצון מהקצבת הזמן בנפרד לכל שאלה בבחינת מפע"ם.....32
איור 8 – התפלגות שביעות הרצון מהעובדה שלא ניתן לחזור לשאלות קודמות בבחינת מפע"ם.....33
איור 9 – התפלגות העדפת הבחינה הממוחשבת על פני בחינת נייר ועיפרון.....33
איור 10– התפלגות הערכת ההצלחה בבחינת הניסוי לעומת הבחינה האמיתית.....34
איור 11 – השוואה בין קבוצת הניסוי לקבוצת הביקורת בהעדפת אחת הבחינות (חשיבה מילולית).....36
איור 12 – השוואה בין קבוצת הניסוי לקבוצת הביקורת בהעדפת אחת הבחינות (חשיבה כמותית).....36
איור 13 – השוואה בין קבוצת הניסוי לקבוצת הביקורת בהעדפת אחת הבחינות (אנגלית).....37
איור 14 – השוואה בין קבוצת הניסוי לקבוצת הביקורת בהעדפת אופן סימון התשובות לשאלות.....38
איור 15 – השוואה בין קבוצת הניסוי לקבוצת הביקורת בהערכת רמת הריכוז בבחינה.....38
איור 16 - השוואה בין קבוצת הניסוי לקבוצת הביקורת בהערכת רמת החרדה בבחינה.....39
איור 17 – השוואה בין קבוצת הניסוי לקבוצת הביקורת בהעדפת מספר השאלות בבחינה.....39
איור 18 – השוואה בין קבוצת הניסוי לקבוצת הביקורת בשביעות הרצון מזמן הבחינה הכולל.....40

תמצית

מחקר זה בחן את היעילות והתקפות של בחינה פסיכומטרית אדפטיבית ממוחשבת (מפע״ם) כהתאמה לנבחנים בעלי הפרעת קשב וריכוז (ADHD). בהתבסס על השערת השיפור הדיפרנציאלי הניח המחקר כי מציאת שיפור דיפרנציאלי בציוני נבחנים בעלי ADHD לעומת ציוני נבחנים ללא לקויות תתמוך בתוקף מתן הבחינה כהתאמה לנבחנים אלו. הרציונל להשערה מבוסס על ההנחה ששיפור דיפרנציאלי כזה משקף מצב בו ההתאמה מפצה על מאפיין שהוא ייחודי ל-ADHD.

במחקר השתתפו 258 נבחנים, מהם 119 בעלי ADHD ו־139 ללא ADHD. הנבחנים בעלי ADHD היו נבחנים שפנו ליחידה לבחינות מותאמות בבקשה להתאמת תנאי הבחינה. הנבחנים חולקו לקבוצות באופן אקראי, והקבוצות נבחנו באחת משתי בחינות ניסוי – בחינת נייר ועיפרון ובחינה אדפטיבית ממוחשבת (מפע״ם).

הממצאים לא הצביעו על שיפור דיפרנציאלי בציוני קבוצת הנבחנים בעלי ADHD שנבחנו בבחינה אדפטיבית ממוחשבת. כמו כן, לא נמצאו הבדלים בין הקבוצות מבחינת מדדי הטעות של הבחינה. נמצא כי נבחנים בעלי ADHD השתמשו בזמן ארוך יותר כדי לפתור את השאלות בכל אחד מתחומי הבחינה, למעט בשאלות הבנת הנקרא באנגלית. עוד נמצא כי קוראות מומחיות, אשר העריכו באופן בלתי תלוי באיזו מידה לדעתן ישתפר ציונו של הנבחן בבחינת מפע״ם לעומת ציונו בבחינת נייר ועיפרון – וזאת על סמך נתוני איטיותו התפקודית, מוסחוּתו וקשייו בארגון (נתונים הלקוחים מתיקי הנבחנים ביחידת במו״ת) – הצליחו להבחין בין נבחנים בעלי ADHD שציונם בבחינת מפע״ם עלה על ציונם בבחינת נייר ועיפרון לבין נבחנים בעלי אבחנת ADHD שציונם לא עלה על ציונם בבחינת נייר ועיפרון. המחקר דן בהשלכותיהם התאורטיות והמעשיות של הממצאים על קביעת ההתאמות.

מבוא

הפרעת קשב וריכוז (ADHD) היא אחת ההפרעות ההתפתחותיות המאובחנות ביותר בגיל הילדות – ממחקרים עולה ששכיחות אבחנת הפרעה זו היא 3%–10% מבני גיל בית הספר (Barkley, 2007). ADHD מוגדרת כהפרעה נוירו־התפתחותית המתאפיינת בקושי לשמור על קשב, בהיפראקטיביות ובאימפולסיביות, וכן במגוון רחב של תסמינים התנהגותיים, קוגניטיביים ורגשיים. ההפרעה מאובחנת אם כמה מהתסמינים מופיעים לפני גיל 12 ובשתי סביבות שונות, תוך פגיעה בתפקודים חברתיים, אקדמיים או מקצועיים (Diagnostic and Statistical Manual of Mental Disorders, (5thed.); DSM-5; American Psychiatric Association, 2013). מחקרי אורך מראים ש־50%–80% מהילדים שאובחנה אצלם ההפרעה ימשיכו לסבול מתסמיניה גם בבגרותם, אם כי יתכן שכמה מהתסמינים ייעלמו או יתמתנו (Barkley, 2007).

הבסיס הקוגניטיבי של ADHD

חוקרים רבים ניסו לזהות את הליקויים הקוגניטיביים והנוירופסיכולוגים העומדים בבסיס ההפרעה, אך עדיין אין הסכמה לגבי סוגיה זו. מחקרים בתחום זה שנעשו בילדים מצביעים על טווח ליקויים נרחב. ברוב המחקרים התחומים שנמצאו במוקד הלקות הם קשב מתמשך, תפקודים ניהוליים ועיכוב תגובה (ראו למשל Barkley, 2007).

אחת התאוריות העיקריות בתחום היא התאוריה של Barkley (1998, 2007), שלפיה מוקד ההפרעה הוא הכשל בעיכוב תגובה. כשל זה גורם לליקוי בתפקודים הניהוליים – האופרציות הקוגניטיביות הגבוהות המופקדות על התנהגות מוכוונת מטרה כגון יזימה, תכנון, התארגנות, הפעלת זיכרון עבודה, בקרה ועוד.

תאוריות אחרות גורסות שמקור ההפרעה בכמה תפקודים בסיסיים ולא בתפקוד יחיד (למשל Tsal, Shalev & Mevorach, 2005; Wåhlstedt, Thorell & Bohlin, 2009). לדוגמה, התאוריה של Tsal et al. (2005) גורסת שיש ארבע פונקציות קשב נפרדות זו מזו– קשב מתמשך, קשב סלקטיבי, הכוונת קשב וקשב ניהולי – ובמקרים של ADHD אחת מהן או יותר נפגעת.

השפעת ADHD על תפקוד קוגניטיבי, אקדמי ויומיומי

מלבד השאלה מהו המקור להפרעה עולה עוד שאלה חשובה: מהם התפקודים הקוגניטיביים הפגועים אצל בעלי ADHD. המחקר כיום מצביע על מגוון רחב של תפקודים קוגניטיביים כאלו, ובהם זיכרון מילולי, קשב ממוקד, קשב מתמשך ופתרון בעיות מורכבות הדורשות שימוש בזיכרון עבודה (Schoechlin & Engel, 2005) ומהירות עיבוד מידע (Rucklidge & Tannock, 2002; Holdnack, Moberg, Arnold, Gur & Gur, 1995).

נוסף על כך, לעתים קרובות ה־ADHD משפיעה על תפקודם האקדמי והתעסוקתי ועל התפתחותם הרגשית־חברתית של הלוקים בה. במחקרים על מבוגרים בעלי ADHD נמצא שיעור גבוה של חזרה על שנות לימוד, הסתייעות בשיעורי עזר, לימוד בכיתות מיוחדות, תת־הישגיות

אקדמית וקשיי קריאה (Biederman, Faraone, Spencer & Wilens, 1993; Breslau et al., 2009; Frazier, Youngstrom, Glutting & Watkings, 2007; Galera, Melchior, Chastang, Bouvard & Fombonne, 2009).

לסיכום, ADHD היא הפרעה התפתחותית המשפיעה על תפקודים קוגניטיביים רבים ומגוונים, ויש עדויות להשפעתה גם על תפקודים אקדמיים ובפרט על יכולת ההיבחנות. יותר ויותר נבחנים בעלי ADHD פונים בבקשה לקבל התאמות בלמידה ובדרכי היבחנות – הן במסגרת לימודיהם התיכוניים והן במסגרות על תיכוניות (מסגרות אקדמיות ומסגרות הכשרה). על כן חשוב לבחון דרכי פיצוי שיאפשרו להם להיבחן באופן תקף והוגן מבחינתם ומבחינת כלל אוכלוסיית הנבחנים.

מבחנים גורליים – הגדרה ומאפיינים עיקריים

בעשורים האחרונים גבר השימוש במבחנים סטנדרטיים גורליים (High-Stakes Tests) במערכת החינוך. מדובר במבחנים בעלי חשיבות גבוהה, אשר עשויים להשפיע על גורלם של נבחנים, צוותים ומוסדות חינוכיים (Amrein & Berliner, 2002). את המונח "מבחן גורלי" (High-Stakes Test) מקובל להגדיר כמבחן שתוצאותיו עשויות להשפיע על אדם או על גוף כלשהם במידה רבה ולגרום להם לרווח או להפסד. המבחן נחשב גורלי יותר ככל שהסיכון הכרוך בו לנבחן רב יותר. המבחנים הגורליים עשויים לקבוע באופן ישיר את דרכו האקדמית של כל תלמיד, ועל כן יש דיון ער בצורך להעניק לכל חלקי האוכלוסייה הזדמנות שווה להשכלה ולהצלחה בהם.

כיום מוסדות אקדמיים רבים בעולם עורכים מבחנים סטנדרטיים וגורליים כדי למיין מועמדים ולדרגם. למבחנים אלו השפעה מכרעת על סיכויי הקבלה ללימודים לתואר ראשון ולעתים אף לתארים מתקדמים. המבחנים הללו משמשים כלי לחיזוי הצלחת המועמדים בלימודים הגבוהים ולדירוגם, וכל מוסד אקדמי רשאי להחליט איזו חשיבות הוא מייחס לציוניהם. לאור המשקל הרב שיש למבחנים אלו ולהשפעתם האפשרית על עתידו של הנבחן, חשוב להתאים את תנאי הבחינה לנבחנים בעלי לקויות מגוונות, ובכללן ADHD, כך שהמדידה תישאר תקפה והוגנת.

מבחנים מואצים לעומת מבחני עוצמה

מבחנים סטנדרטיים גורליים במערכת החינוך יכולים להיות "מבחני עוצמה" (Power Tests) או "מבחנים מואצים" (Speed Tests). מבחני עוצמה אינם מתיימרים לבחון את מהירות העבודה של הנבחן, אלא את רמת הידע או היכולת שלו, ועל כן הגבלת זמן בהם אינה הכרחית והיא אף עלולה לפגוע בתוקף תוצאותיהם. מבחנים מואצים, לעומת זאת, נועדו בין השאר להעריך מהי מהירות העבודה של הנבחן, ועל כן הגבלת הזמן בהם נחוצה.

בשל אילוצים לוגיסטיים ומנהלתיים בדרך כלל אין ברירה אלא להגביל את זמן המבחן, ולכן גם במבחני עוצמה יכולה להופיע הגבלת זמן כלשהי. כך, למשל, מבחני הקבלה לאוניברסיטאות (כמו הבחינה הפסיכומטרית) הם מואצים חלקית (קנת-כהן, ברונר ואיתן, 2008).

הגבלה זו עלולה להקשות במיוחד על נבחנים שתפקודם אטי בשל ליקוי או מגבלה אשר אינם אמורים להימדד ישירות בבחינה. למשל, נבחנים בעלי לקות קריאה (דיסלקסיה) נוטים לקרוא

לאט יותר מנבחנים שאינם סובלים מלקות זו, ולכן תפקודם בבחינה כתובה עלול להיות אטי יותר ללא קשר לרמת בקיאותם בחומר. יש לתת את הדעת לתנאי ההיבחנות של אוכלוסיות אלו מאחר שגורם ההאצה בבחינה עלול לגרום לשתי בעיות: א. פגיעה במדידה, מאחר שהתוצאה מושפעת ממשתנים שאינם רלוונטיים ליכולת הנמדדת; ב. פגיעה בהוגנות הבחינה כלפי נבחנים בעלי לקויות, המביאה לידי פגיעה בעקרון שוויון ההזדמנויות.

הבחינה הפסיכומטרית בישראל

הבחינה הפסיכומטרית בישראל שעורך המרכז הארצי לבחינות ולהערכה (מאל"ו) משמשת למיון מועמדים ללימודי תואר ראשון ברבים ממוסדות ההשכלה הגבוהה בישראל (ובכללם כל המוסדות האוניברסיטאיים), והיא מבקשת לחזות את סיכויי ההצלחה של סטודנטים בלימודי התואר הראשון. בבחינה זו שאלות משלושה תחומים: חשיבה מילולית, חשיבה כמותית ואנגלית. בשנתיים האחרונות נוספה לשאלות בתחום החשיבה המילולית גם מטלת כתיבה. כל מוסד אקדמי רשאי להחליט איזה משקל ייתן לציוני הבחינה בקבלה ללימודים לעומת ציוני בחינות הבגרות. במקרים רבים ניתן לציוני הבחינה הפסיכומטרית משקל רב בהחלטה, ועל כן היא נחשבת בחינה גורלית (אנגל, 2008).

הבחינה הפסיכומטרית נחשבת בחינה אובייקטיבית, סטנדרטית ובעלת תוקף חיזוי טוב להצלחה בלימודי התואר הראשון. תוקף הניבוי של ציון הסֶכֶם (הכולל את ציון הבחינה הפסיכומטרית ואת ציוני הבגרות) בניבוי ממוצע הציונים בסוף שנה א' באוניברסיטאות עומד על 0.50, והוא יציב לאורך שנים (אורן, קנת-כהן וברונר, 2007).

במחקרים שנערכו בנוגע להקצאת הזמן בבחינה הפסיכומטרית נמצא שתוספת זמן משפרת בממוצע את ציוניהם של כל הנבחנים, ואילו בבדיקת דירוג הנבחנים שנבחנו עם תוספת זמן ובלעדיה, במסגרת בדיקת הוגנות, נמצאו ממצאים מעורבים (קנת-כהן, ברונר ואיתן, 2008). ממצאים אלה מעידים שהבחינה מואצת, לפחות במידה כלשהי.

לאור זאת, אין זה מפתיע כי תוספת זמן היא אחת ההתאמות המבוקשות ביותר בבחינה הפסיכומטרית, בעיקר בקרב נבחנים אשר חשים כי מגבלה כלשהי (מגבלה רפואית, מגבלה נפשית או לקות למידה) מאטה את תפקודם ופוגעת בביצועיהם. עם זאת, ההחלטה על מתן תוספת זמן בבחינה הפסיכומטרית היא החלטה סבוכה, ובעת קבלתה יש להתחשב בשיקולים פסיכומטריים של שמירה על איכות המדידה ותקפותה ובשיקולים של הוגנות ואחריות הנוגעים הן לנבחנים בעלי לקויות והן לכלל הנבחנים. שיקולים אלו, העוסקים בהטיות במדידה שתוספת זמן בבחינה עלולה לגרום, יידונו בהרחבה להלן.

התאמות במבחנים גורליים – משמעות והשלכות

במבחני הישגים סטנדרטיים נהוג להניח כי עיקר ההבדל בציונים בין נבחנים נובע משוני ביניהם ביכולת הנמדדת ולא משוני בנוסח המבחן או באופן עריכתו. עם זאת, מבחנים סטנדרטיים רבים אינם הוגנים כלפי תלמידים בעלי לקויות, משום שהם מתקשים להפגין בהם את יכולותיהם ואת הידע שלהם, שלא כמו שאר התלמידים (Sireci, Scarpati, & Li, 2005). מבחינה תאורטית לקויות אלו אינן אמורות להשפיע על היכולת המהותית שהבחינה מבקשת למדוד, אך מבחינה מעשית הן עלולות לפגוע קשות ביכולתו של הנבחן לבטא אותה, וזאת בשל אופן ההעברה של

הבחינה. למשל, אם עיוור ייבחן בבחינה סטנדרטית בהיסטוריה בלי שיותאמו עבורו אופן הצגת השאלות והמענה עליהן, סביר שיתקשה לבטא את הידע שלו גם אם הוא שולט בחומר היטב. התאמות במבחנים נועדו להתמודד עם קושי זה.

התאמות בבחינות הן שינויים מכוונים בהליך המבחן שנועדו לעקוף מגבלות שאינן רלוונטיות ליכולת שהוא מודד וכך להנגיש אותו לתלמידים בעלי לקות ולשפר את התוקף ואת הדיוק של מדידת ביצועיהם (Sireci et al., 2005). מבחינה ערכית וחברתית נועדו ההתאמות לאפשר לכל אדם להגיע להישגים התואמים את יכולתו ולממש את הפוטנציאל שלו, וכך לקדם את עקרון שוויון ההזדמנויות. מבחינה פסיכומטרית נועדו ההתאמות לשפר את התוקף וההוגנות של הציונים בבחינות ולהקטין את טעויות המדידה בהן. ביסוד מתן התאמות עומדת ההנחה שעל הציונים לשקף את יכולות הנבחן בלי שלקויות שאינן רלוונטיות לבחינה ישפיעו עליהם ובלי שיהיה צורך לתקנם (Cohen, Ben-Simon, Moshinsky & Eitan, 2008).

בעבר ניתנו התאמות בעיקר לתלמידים בעלי לקויות פיזיות או חושיות, אך בעשורים האחרונים הורחב השימוש בהתאמות, וכיום הן ניתנות גם ללקויי למידה (קלפר, טורוול, קנת-כהן ואורן, 2013). כיום מתן התאמות הוא הדרך הנפוצה ביותר להתמודדות עם בחינת תלמידים בעלי לקויות, וזאת במבחנים סטנדרטיים בכלל ובמבחנים גורליים בפרט.

סוגי ההתאמות רבים ומגוונים, והם כוללים שינויים בדרך הצגת המבחן (כתב מוגדל, הקראה), שינויים בדרך מתן התשובות (תשובות בעל פה, כתיבה במחשב), שינויים בתנאי הבחינה (חדר מבודד), שינויים הקשורים למשך הבחינה (הארכת זמן, מתן הפסקות) והתרת שימוש בעזרי בחינה (מחשב כיס, מילון או מילונית אלקטרונית). בעת מתן התאמות יש לקחת בחשבון את הנבחן ואת הלקות שלו מצד אחד, ואת אופייה של הבחינה ואת הכישורים שהיא בודקת מצד שני (Cohen et al. 2008). ההתאמה השכיחה ביותר בבחינות היא תוספת זמן, הן לבדה והן בלוויית התאמות אחרות (כגון השמעה) (Thurlow, Lazarus, Thompson & Morse, 2005).

בדיקת תוקפן ויעילותן של התאמות – השערת האינטראקציה

קשה להתאים תנאי בחינה בלי לפגוע בעקרונות הפסיכומטריים הקובעים שהמדידה צריכה להיות סטנדרטית ואחידה לכל הנבחנים. ראשית, קשה לשנות רק את המרכיבים שאינם רלוונטיים למדידת היכולת הנבדקת אך לשמור על המרכיבים המהותיים לה. שנית, קשה להשוות בין הציונים במבחן עם התאמות לציוני מבחן ללא התאמות, משום ששני המבחנים שונים זה מזה. נוסף על כך, בספרות המקצועית אין תמימות דעים בשאלה אם ההתאמות הוגנות כלפי התלמידים בכלל וכלפי תלמידים בעלי לקויות למידה בפרט. למשל, אם כל הנבחנים יכולים להרוויח מהתאמה כלשהי, אך רק מעטים מהם מקבלים אותה, עלול להיווצר מצב שאינו הוגן (Phillips, 1994; Sireci et al., 2005).

כדי להתמודד עם בעיות אלו נוסחה "השערת האינטראקציה" (The Interaction Hypothesis), ולפיה יש לתת רק התאמות המביאות לידי שיפור בציוני לקויי למידה בהשוואה לציוניהם במבחן סטנדרטי, אך אינן מביאות לידי שיפור בציוני מי שאינם לקויי למידה (Sireci et al., 2005).

השערה זו מבוססת על ההנחה שאם התאמה מנטרלת הקושי שהלקות גורמת לו, רק נבחנים עם לקות זו יפיקו ממנה תועלת.

במקרים שבהם אין נוהגים לפי השערת האינטראקציה ייתכן שנבחנים ללא התאמות יופלו לרעה, מה שיפגע בהוגנות המבחן ובתוקפם של הציונים. עם זאת, הדרישה לנהוג לפי השערה זו ללא סייג היא דרישה תובענית מאוד, וכמעט שאין התאמות העומדות בה. גרסה מוחלשת של השערת האינטראקציה דורשת שיפור דיפרנציאלי (Differential Boost). לפי גרסה זו יש לתת התאמה כאשר היא מסייעת לתלמידים בעלי לקויות יותר מאשר לשאר האוכלוסייה, אף על פי שגם לתלמידים ללא לקויות היא מסייעת במידה כלשהי (Fuchs & Fuchs, 1999). מחקרים הבודקים התאמות הניתנות בפועל מציגים ממצאים מעורבים בשאלה אם השערת האינטראקציה והשערת השיפור הדיפרנציאלי מתקיימות: כמה מהם משיבים על כך בחיוב, וכמה – בשלילה.

Sireci et al. (2005) ערכו מטא־אנליזה ל־28 מחקרים שבדקו אם השערת האינטראקציה מתקיימת במבחנים מותאמים של תלמידי תיכון ושל סטודנטים. ממצאי רוב המחקרים תמכו בקיומו של שיפור דיפרנציאלי: נמצא שיפור בציון הממוצע של כל הנבחנים, אך השיפור היה גדול יותר בקרב קבוצת לקויי הלמידה. ממצא זה בלט במיוחד בהתאמה של מתן תוספת זמן, אך הופיע גם בהתאמה של הקראת השאלות במבחנים בחשבון. עם זאת, עורכי המחקר הסיקו כי בשל ריבוי סוגי ההתאמות ודרכי העברתן ובשל ההטרוגניות של מקבליהן אי אפשר להכריע באופן גורף בשאלה אם מתן ההתאמות יעיל ואם השערת האינטראקציה תקפה.

מטא־אנליזה אחרת (Gregg & Nelson, 2012) התמקדה במחקרים שבחנו את השפעתה של תוספת הזמן על תפקודם של מתבגרים בעלי לקויות למידה. הממצא העיקרי במחקר זה היה מחסור במחקר כמותי אמין ורלוונטי בתחום. מ־2,300 מחקרים שעסקו בנושא נמצאו רק 9 שעמדו בקריטריונים הבסיסיים שהציבו כותבי המאמר (כגון מדגם גדול מ־9, אבחנה ברורה של לקות למידה ונגישות של הנתונים המספריים), והשונות הרבה באופן הגדרת לקות הלמידה, בסוג המבחן הגורלי, בשאלות הספציפיות שנשאלו הנבחנים ובתוצאות המבחנים הקשתה על ההשוואה בין מחקרים אלו. באופן כללי נמצא כי תוספת זמן משפרת הן את ביצועיהם של לקויי הלמידה והן את ביצועי חברי קבוצת הביקורת, וכן נמצא כי ביצועיה הממוצעים של קבוצת הביקורת בתנאי בחינה רגילים עדיין עולים על ביצועיה של קבוצת לקויי הלמידה המקבלים תוספת זמן. מטא־אנליזה זו הראתה כי אין די נתונים כדי להכריע בשאלה אם יש אינטראקציה בין רמת השיפור לסוג הקבוצה, כפי שהשערת האינטראקציה תובעת.

Zuriff (2000) ערך חמישה מחקרים בניסיון לבדוק אם השערת האינטראקציה תקפה כאשר נותנים התאמה של תוספת זמן. הוא מצא כי בדרך כלל התאמה זו מסייעת הן לתלמידים לקויי למידה והן לתלמידים שאינם לקויי למידה, וקבע שאין בכך כדי לאשש את השערת האינטראקציה. עם זאת, בחלק מהמחקרים נעשה שימוש במבחנים מואצים (כמו מבחן קריאה Nelson-Denny), שבהם צפוי שכל הנבחנים ישפרו את הישגיהם אם יקבלו תוספת זמן. לעומת זאת, במבחנים לא מואצים, כמו ה־ACT (הדומה ל־SAT), תמכו הממצאים בהשערת האינטראקציה.

גישה אחרת לנושא (Cahan, Alkoby & Nirel, 2010) טוענת כי מתן התאמות ללקויי למידה בלבד הוא בעייתי, מכיוון שלקות הלמידה אינה הגורם העיקרי המשפיע על מידת השיפור בהישגי הנבחן בבחינה עם התאמות. אם יימצא שהשערת האינטראקציה אינה תקפה לגבי התאמה כלשהי, סביר להניח כי הלקות אינה הגורם המרכזי המשפיע על יכולת להפיק תועלת מההתאמה זו. מכאן עולה שמתן התאמות ללקויי למידה פוגע ברוב הנבחנים, שאינם לקויי למידה, ולכן הוא פוגע בהוגנות המבחן ובתוקפו.

ביצוע נבחנים בעלי ADHD במבחנים גורליים – השפעתה של תוספת זמן

בארץ ובעולם מקובל מאוד להתאים בחינות לנבחנים בעלי ADHD (הפרעת קשב וריכוז) באמצעות תוספת זמן. כמו במקרה של לקויות למידה, גם כאן עולות שאלות בנוגע ליעילות ההתאמות, להוגנותן ולקיום השערת האינטראקציה. עם זאת נערכו רק מחקרים מועטים העוסקים בשאלות אלו, ואין מסקנה אחידה וברורה העולה מתוצאותיהם (Pariseau, Fabiano, Massetti, Hart & Pelham, 2010).

Pariseau ואחרים (2010) בחנו את תפקודם של ילדים בני 9–12 בעלי הפרעת קשב וריכוז במטלות קריאה, כתיבה וחשבון, עם תוספת זמן ובלעדיה. מצד אחד, נמצא כי הקבוצה שקיבלה תוספת זמן ענתה נכונה על יותר שאלות מן הקבוצה שנבחנה ללא תוספת זמן. מצד שני, הספק ההצלחה (מספר התשובות הנכונות לדקה) של הקבוצה שקיבלה תוספת זמן היה נמוך משמעותית מזה של הקבוצה האחרת. נוסף על כך, נמצא כי אחוזי הדיוק בתשובות אצל תלמידים שקיבלו תוספת זמן היו גבוהים מעט יותר (באופן מובהק) מאשר אצל תלמידים שלא קיבלו אותה.

מחקר אחר (Lewandowski, Lovett, Parolin, Gordon & Codding, 2007) בדק את תפקודם של תלמידים בגיל חטיבת הביניים, בעלי ADHD ובלעדיה, בבחינה במתמטיקה. התוצאות לא תמכו בהשערת האינטראקציה: קבוצת בעלי ה־ADHD לא הפיקה מתוספת זמן של 50% תועלת רבה משהפיקה ממנה קבוצת הביקורת. כמו כן נמצא כי קבוצת בעלי ה־ADHD מתאפיינת בתפקוד נמוך מזה של קבוצת הביקורת מבחינת מהירות עיבוד מידע, קצב ודיוק. במחקר של Brown, Reichel and Quinlan (2011), שהתמקד בהבנת הנקרא, נמצא כי תוספת זמן הביאה לידי שיפור ניכר בביצועיהם של מתבגרים בעלי ADHD. ההסבר שניתן לשיפור התבסס על התפקוד הנמוך מהנורמה של קבוצת בעלי ה-ADHD במטלות סטנדרטיות המצריכות שימוש בזיכרון עבודה ועיבוד מידע מהיר. מכיוון שתפקודם של חברי קבוצה זו במטלות הבנת שפה כלליות לא נמצא נמוך במובהק מהנורמה, נטען כי תוספת זמן מפצה על העיבוד האטי ועל ההנמכה בזיכרון העבודה ומביאה לידי מדידה מדויקת יותר של הבנת הנקרא שלהם. עם זאת, מאחר שלא נערכה השוואה בין תפקוד קבוצה זו לתפקוד קבוצה של נבחנים ללא ADHD, אין לדעת אם שיפור זה ייחודי לקבוצת בעלי ה־ADHD או נובע מארטיפקטים סטטיסטיים, כגון רגרסיה לממוצע.

בשנים האחרונות חלה עלייה חדה במספר הבקשות לתנאים מותאמים בבחינה הפסיכומטרית בגין ADHD. מנתוני מאל"ו עולה כי בשנים 2002–2012 חלה עלייה של 258% במספר הפניות הכללי לקבלת תנאים מותאמים, עלייה של 310% במספר הפניות בגין לקויות למידה ועלייה של 590% במספר הפניות בגין הפרעת קשב וריכוז. לצורך השוואה, העלייה הכללית במספר הנבחנים בבחינה הפסיכומטרית בתקופה זו הייתה כ־29%. עלייה זו מציבה אתגר של ממש בפני האחראים על קביעת התנאים לפונים. בתקופה זו חלה עלייה גם בשיעור האבחנה של ADHD, והיא לוותה בהתרופפות רבה בקריטריונים לאבחון ההפרעה, הן בארץ והן בעולם. האבחנה ניתנת בקלות רבה יותר בהשפעת אינטרסים של הרופאים, המורים, ההורים והמאובחנים עצמם, וכן בהשפעת אינטרסים מסחריים של חברות תרופות גדולות (Timini, 2008).

יותר ויותר אנשי מקצוע מתייחסים להתנהגויות נורמטיביות כעשויות לשקף הפרעת קשב וריכוז, מה שמקשה את אבחונה. האבחנה קשה אף יותר אם היא נעשית בגיל הבגרות ומסתמכת ברובה על דיווח עצמי של הפונה, ללא שימוש בקריטריונים "קשים". מאחר שכיום קיום ההפרעה נקבע בעיקר על סמך דיווח עצמי על התסמינים ועל שכיחותם, הגבול בין תפקוד תקין לבין תפקוד לקוי בעקבות הפרעה אינו חד. רוב המומחים בתחום מסכימים כי התסמינים פתוחים לפרשנות של הפונה, של הוריו, של מוריו ושל רופאיו. אכן, שיעור הפניות של מורים והורים החושדים בקיומה של הפרעת קשב וריכוז אצל תלמיד עולה באופן מתמיד, ולעתים קרובות פניות אלו אינן תואמות למצבו של האדם בפועל. ההערכה המקובלת היא שעד 10% מן האוכלוסייה לוקים בהפרעת קשב וריכוז, אך כיום שיעור אבחנת ההפרעה בילדים עד גיל 18 בארה"ב עומד על כ־15% (Schwartz, 2013, December, 14).

כמו כן קיימת שונות רבה באיכות האבחונים, והיא נובעת משונות במידת ההיכרות של המאבחן עם הפונה (למשל, פגישה אחת עם רופא לעומת היכרות טיפולית רבת שנים), משונות בסוג הבדיקות הנערכות ובאיכותן (דיווח עצמי או מילוי שאלון לעומת איסוף נרחב של הרקע ההתפתחותי והאקדמי ושימוש במבחני קשב מתוקפים) ומשונות במידת הבקיאות של המאבחן בשלילת אבחנות מבדילות, כלומר אטיולוגיות נפשיות, פיזיות וסביבתיות חלופיות. לכמה מהמאבחנים אף חסרה ההכשרה המתאימה לביצוע מטלות אלו.

סוגיה נוספת שראוי להזכיר היא תופעת ההתחלות (malingering). לטענת Jasinski and Ranseen (2011) בעיה זו רלוונטית במיוחד כשמדובר באבחון הפרעת קשב וריכוז, הן משום שתהליך האבחון מתבסס במידה רבה על דיווח עצמי והן משום שמי שאובחנה אצלו ADHD עשוי להפיק מכך תועלת אישית: לקבל הטבות כספיות (פיצויים, אחוזי נכות וכדומה), לקבל תרופות מרשם מבוקשות ולקבל התאמות במבחנים גורליים.

Harrison, Edwards and Parker (2007) ערכו ניסוי ובו ביקשו מקבוצת סטודנטים להעמיד פנים שיש להם ADHD, הן בעת שאלוני דיווח עצמי של תסמיני ADHD והן בביצוע מטלות הבוחנות שטף קריאה ומהירות עיבוד מידע. ביצועי קבוצה זו של "מתחזים" הושוו לביצועיה של קבוצת סטודנטים בעלי ADHD ולביצועיה של קבוצת ביקורת. נמצא כי קל מאוד לחקות תסמינים של

ADHD באופן שיזכה את המִתְחַלה באבחנה הרצויה, וזאת מכיוון שכלי האבחון הקיימים אינם מיטיבים להבחין בין ADHD להִתחלוּת.

לבסוף, לסוגיות האבחנתיות המורכבות מצטרף גם עיוות כלכלי המקל על אנשים מאוכלוסיות מבוססות להיות מאובחנים כבעלי ADHD. אף שמעריכים ששיעור בעלי ADHD בארץ נמוך מ־10%, יש בתי ספר שבהם 60% מהילדים מאובחנים כבעלי ADHD. במוסדות להשכלה גבוהה עד 50% מהסטודנטים מאובחנים כבעלי ADHD או נוטלים תרופות לטיפול בתסמיני הפרעה זו. מאחר שאבחנה של ADHD מזכה תלמידים בהתאמות, תופעת אבחון היתר הולכת וגוברת, וזאת אף על פי שרבים מהמאובחנים כבעלי ADHD אינם עומדים אפילו בקריטריונים הגמישים ביותר (לוינסקי, 2014).

בכל הנוגע לבחינה הפסיכומטרית, אבחון יתר של ADHD צפוי לפגוע בהוגנות הבחינה כלפי כלל הנבחנים. תוספת זמן בבחינה, שכאמור היא אחת ההתאמות המבוקשות ביותר, משנה את רמת הביצוע בבחינה במידה רבה, ולכן מתן בלתי מבוקר שלה עלול לפגוע בהוגנות הבחינה.

כיום נבחנים בעלי ADHD אינם מקבלים תוספת זמן בבחינה הפסיכומטרית באופן אוטומטי, וזאת הן בשל הפגיעה בהוגנות הכרוכה בכך והן משום שלא נמצאה עדות חד משמעית לכך שתוספת זמן אמנם משפרת דיפרנציאלית את תפקודם (אנגל, 2008). ההתאמות המקובלות לנבחנים אלו בבחינה הפסיכומטרית הן יציאה להפסקות בנות 5 דקות בין פרקי הבחינה (כל 40 דקות) לשם התרעננות ומנוחה, והיבחנות בחדר מעוט נבחנים. התאמות נוספות שנבחנים בעלי ADHD עשויים לקבל בהתאם לצורך הן שימוש במחשבון בפרקים הכמותיים, שימוש בגיליון תשובות מוגדל ותוספת זמן במטלת החיבור. התאמות אלו ניתנות כדי לפצות על אימפולסיביות וקשיי ארגון המאפיינים את ההפרעה.

מבחן אדפטיבי ממוחשב – מפע"ם

בשל הבעיות המהותיות הנובעות ממתן תוספת זמן במבחנים גורליים מואצים פיתח מאל״ו מבחן ממוחשב אדפטיבי. מבחן זה הוא מבחן עוצמה, שבו רמת קושי הפריטים קובעת את הציון ולא הספק הנבחן. המבחן פותח כדי שיהיה אפשר לפצות נבחנים שנקבע כי הם אמנם זקוקים לתוספת זמן, וזאת בלי לסבול מעיוותי מדידה. בחינה זו, המכונה מפע״ם (מבחן פסיכומטרי על גבי מחשב), משמשת כיום לפיצוי נבחנים בעלי לקויות למידה מוכחות ומתועדות הזקוקים לתוספת זמן, וכן נבחנים בעלי מגבלות פיזיות מגוונות, הנבחנים בשפה העברית. הבחינה נערכת על גבי מחשב, ובכל שלב בה האלגוריתם הממוחשב בוחר עבור הנבחן שאלה שרמת הקושי שלה נקבעת לפי הישגיו בשאלות הקודמות (ראו פירוט לגבי פיתוח הבחינה ואופן קביעת הציון בפרק שיטת המחקר. אצל נבחנים ללא מגבלות נמצא שהבחינה שקולה לבחינה בנייר ובעיפרון, ואילו אצל נבחנים בעלי לקויות למידה ומגבלות פיזיות נמצאו שיפורים דיפרנציאליים (Moshinskey & Kazin, 2004). מסיבות רבות ומגוונות האפשרות להיבחן במפע״ם אינה פתוחה לכל הנבחנים, ובחינה זו אינה ניתנת כיום לנבחנים בעלי ADHD. מטרת מחקר זה הייתה לבדוק אם נבחנים בעלי ADHD משפרים את ציונם באופן דיפרנציאלי כשניתנת להם התאמה זו, וזאת כדי ללמוד אם כדאי לתתה להם.

מטרות המחקר

למחקר זה כמה מטרות:

1. להשוות את תפקודם של נבחנים המאובחנים כבעלי ADHD בבחינת מפע״ם לתפקודם בבחינת נייר ועיפרון (ני״ע) מותאמת ולתפקודם של נבחנים ללא ADHD בשני סוגי הבחינות. כפי שצויין, אצל נבחנים ללא ADHD בחינת נייר ועיפרון שקולה לבחינת מפע״ם, כלומר אין הפרש ניכר בין ממוצעי ציוני הקבוצות בבחינות אלו. לפיכך, אם יימצא בקבוצת הנבחנים המאובחנים כבעלי ADHD פער משמעותי בין ממוצעי הציונים בשתי הבחינות לטובת בחינת המפע״ם, אפשר יהיה להסיק מכך כי בחינה זו יכולה לשמש התאמה הולמת לנבחנים בעלי ADHD.

2. לבחון את ההשערה שטיב ההתכנסות של ציוני בחינות מפע״ם של נבחנים בעלי ADHD נמוך מזה של ציוני בחינות של נבחנים ללא ADHD. טיב ההתכנסות הוא מדד המושפע מתנודתיות בתפקוד הנבחן. מאבחני ADHD מרבים לטעון שתפקודם של מאובחניהם בבחינות מתאפיין בתנודתיות רבה, כלומר שבעלי ADHD נוטים להשיב תשובות שגויות על שאלות המצויות בתחום הידע שלהם יותר מאחרים, וזאת בשל חוסר יציבות בתפקודם הנובע מירידות זמניות בקשב. מטענה זו עולה כי נבחנים בעלי ADHD נוטים להיכשל בשאלות קלות אך נוטים להצליח בשאלות קשות יותר. לכן ייתכן שהציון של נבחנים אלו בבחינה הרגילה נפגע בשל טעויות בשאלות קלות יחסית, טעויות שאינן משקפות חוסר יכולת (שהרי הם עונים נכונה על שאלות קשות) ונובעות מבעיית הקשב. אם טענה זו נכונה, סביר שנבחנים בעלי ADHD יתאפיינו בביצוע לא אחיד ובמדדי טיב התכנסות נמוכים יותר מנבחנים מאוכלוסיות אחרות, וכן שאחוז גבוה יותר של ציוניהם לא יגיע לכלל התכנסות כלל. ההשערה שנבדקה היא שהביצוע הלא אחיד (המשוער) של בעלי ADHD יגרום לחוסר מונוטוניות בהשתנות טעות המדידה, כלומר למקרים שבהם טעות המדידה גדלה משאלה לשאלה.

3. לבדוק אם נבחנים בעלי ADHD משתמשים בזמן רב יותר מנבחנים ללא ADHD כדי לענות על השאלות לסוגיהן.

4. לברר אם אפשר לנבא אם ציונו של מבקש התאמות בעל ADHD בבחינת מפע״ם יעלה על ציונו בבחינת נייר ועיפרון. אוכלוסיית האנשים שהם בעלי ADHD היא הטרוגנית ביותר, ולכן יש לבדוק לא רק את הציונים הממוצעים הקבוצתיים אלא גם את השונות בתוך הקבוצה, כלומר את תפקודו של כל אחד ואחד ממשתתפיה בבחינת מפע״ם. לשם כך נערך ניתוח אידיוגרפי שהשווה את ממוצע השיפור של כל נבחן לשיפור האופייני הממוצע בקבוצת הנבחנים ללא ADHD, ונבדקה תקפותם של שיפוטי מומחים לגבי השימוש בבחינת מפע״ם כהתאמה שעשויה להניב שיפור בציונו של כל נבחן.

5. לקבל משוב בנוגע לחוויית נבחנים עם ADHD בבחינת מפע״ם. בפרט, נבחנים שנבחנו בבחינת המפע״ם דיווחו בשאלון משוב על חוויית הבחינה, על מידת המוטיבציה שלהם, על

רמת שביעות רצונם מהבחינה ועל הערכתם לגבי תפקודם בבחינה זו לעומת בחינת ני״ע. נתונים אלו הושוו לתפקודם בפועל ולרמת השיפור שהשיגו, אם השיגו, בבחינת מפע״ם לעומת בחינת ני״ע.

הנחות העבודה

1. לגבי נבחנים ללא מגבלות כלשהן בחינת המפע״ם שקולה לבחינה פסיכומטרית רגילה (Moshinsky & Kazin, 2004).

2. פער בין ההישגים בבחינת המחקר (מפע״ם או בחינת ני״ע) להישגים בבחינה האמתית צפוי אצל כל הנבחנים. הציון בבחינה האמתית צפוי להיות גבוה מהציון בבחינת המחקר. לשיפור זה שלושה גורמים: (1) עצם החזרה על הבחינה; (2) ההכנה האינטנסיבית שבין ההיבחנות הראשונה לשנייה – בהקשר זה יש להניח שגם משך הזמן בין שתי ההיבחנויות משפיע על השיפור; (3) עלייה במוטיבציה להצלחה בבחינה האמתית. עם זאת, יש לציין כי הלחץ בבחינה האמתית יכול לקזז את השיפור הנובע מגורמים אלו. בהקשר לגורם העלייה במוטיבציה חשוב לציין כי נבחנים בקבוצת המחקר (נבחנים בעלי ADHD בבחינת מפע״ם) הוזמנו לבצע בחינה שאופן העברתה שונה מאופן הבחינה האמתית שבה הם עומדים להיבחן. ייתכן שהשפעתו של גורם זה מקיימת אינטראקציה עם השפעת אופן ההיבחנות. לכן הושוו הנבחנים גם לקבוצת ביקורת נוספת, שכללה נבחנים בעלי ADHD שנבחנו בבחינת מחקר בפורמט ני״ע באותם התנאים שבהם ייבחנו בבחינה האמתית.

על מנת לבקר את כל ההשפעות שהוזכרו לעיל הושוותה קבוצת המחקר (קבוצה 1 – נבחנים בעלי ADHD שנבחנו בבחינת המחקר במפע״ם) לשלוש קבוצות ביקורת:

1. נבחנים ללא ADHD שנבחנו במפע״ם בבחינת המחקר ובבחינת ני״ע בבחינה האמתית.

2. נבחנים בעלי ADHD שנבחנו בבחינת ני״ע מותאמת הן בבחינת המחקר והן בבחינה האמתית.

3. נבחנים ללא ADHD שנבחנו בבחינת ני״ע הן בבחינת המחקר והן בבחינה האמתית.

שיטת המחקר

משתתפים

המדגם: המדגם מנה 258 נבחנים. כל הנבחנים שהשתתפו במחקר היו בני 16–35, טווח המשקף את גילי רוב הנבחנים בבחינה הפסיכומטרית ומייצג נבחנים שלגביהם הבחינה היא גורלית. 52% מהנבחנים היו גברים ו־48% – נשים, וגילם הממוצע היה 22 (ס"ת 2.46). להלן תיאור מפורט של כל קבוצה:

קבוצת הניסוי: קבוצת הניסוי (קבוצה 1) מנתה 58 נבחנים שאובחנו כבעלי ADHD ונבחנו במסגרת המחקר בבחינת מפע"ם. המשתתפים נבחרו ממאגר הנרשמים שפנו בבקשה לבחינה בתנאים מותאמים, וקיום ADHD אצל אותם פונים נקבע על פי קריטריונים אלו: (1) הצגת מסמכים המעידים על קיומה של ההפרעה כבר מגיל צעיר, כגון אבחונים שנערכו בידי אנשי מקצוע מתאימים (נוירולוג, פסיכולוג, פסיכיאטר או מאבחן דידקטי), חוות דעת של מורים או מטפלים שסייעו לנבחן, ותעודות מבית הספר היסודי של הנבחן המעידות על קשיי תפקוד שלו הנובעים מקשיי קשב וריכוז; (2) אבחון עדכני (אבחון משלוש השנים האחרונות שנערך כשגיל הנבחן עלה על 16) המעיד על הימשכות ההפרעה ועל דרגת חומרה מספקת שלה; (3) ראיה לכך שההפרעה עשויה להפריע לתפקוד הנבחן בבחינה.

לא נכללו במחקר נבחנים שיש להם מגבלה נוספת כלשהי, גופנית או נפשית, ובכלל זה לקות למידה. נוסף על כך לא נכללו במחקר נבחנים אשר הגישו ערעור למדור בחינות מותאמות (במו"ת) לגבי התנאים שקיבלו בבחינה, וזאת כדי למנוע מצב של ניגוד עניינים. כמו כן לא נכללו במחקר משתתפים אשר נבחנו בעבר בבחינה הפסיכומטרית או בבחינות ידע באנגלית (בחינות אמי"ר או אמיר"ם), וזאת כדי למנוע מצב שבו כמה מהנבחנים ייבחנו בניסוי בבחינה ראשונה ואילו נבחנים אחרים ייבחנו בו בבחינה חוזרת, מה שיגדיל את השונות בין הנבחנים.

קבוצות הביקורת: המחקר כלל שלוש קבוצות ביקורת: שתי קבוצות ביקורת של נבחנים ללא ADHD, אחת שנבחנה בבחינת ני"ע (קבוצה 2) ואחת שנבחנה בחינת מפע"ם (קבוצה 4) וקבוצת ביקורת אחת של נבחנים בעלי ADHD שנבחנו בבחינת ני"ע מותאמת (קבוצה 3), כמפורט בלוח 1.

נבחנים ללא ADHD נדגמו אקראית מקרב הנרשמים לבחינה הפסיכומטרית שלא פנו למדור במו"ת, והם נבחנו באותם המועדים שבהם התקיימו הבחינות המותאמות. כדי לוודא שנבחנים אלו אינם סובלים מ־ADHD התבקשו גם הם למלא שאלונים לדיווח עצמי לאבחון ההפרעה. נבחני הביקורת בקבוצה 3 (נבחנים בעלי ADHD) אותרו באופן דומה לאופן שבו אותרו נבחנים בעלי ADHD לקבוצת הניסוי, כמפורט לעיל.

כלים

בחינת הניסוי – גרסת נייר ועיפרון: בחינת הני"ע ששימשה בניסוי היא נוסח ישן של הבחינה הפסיכומטרית ששימש כבחינה אמתית במועדים קודמים. בנוסח 8 פרקים: 2 פרקי חשיבה מילולית, 2 פרקי חשיבה כמותית ו־2 פרקי אנגלית. נוסף על כך הנוסח כולל שני פרקי פיילוט. בכל אחד מפרקי החשיבה המילולית 30 שאלות, בכל אחד מפרקי האנגלית 27 שאלות ובכל אחד

מפרקי החשיבה הכמותית 25 שאלות. בחינה זו הועברה לשתי קבוצות הביקורת שנבחנו בבחינת נ״ע: לקבוצה 3, שנבחנה בבחינת נ״ע מותאמת, ולקבוצה 4, שנבחנה בבחינת נ״ע רגילה.

בחינת הניסוי – בחינת מפע״ם: בחינת מפע״ם היא בחינה אדפטיבית ממוחשבת שנכנסה לשימוש ביולי 2000. בבחינה שלושה פרקים: פרק חשיבה מילולית, פרק חשיבה כמותית ופרק אנגלית, המוצגים לנבחנים בסדר זה. הבחינה בנויה כבחינת עוצמה על פי מודל IRT: היא בוחרת שאלות לכל נבחן בהתאם לרמת הביצוע שלו בעת פתרון השאלות הקודמות. הבחינה נפתחת בשתי שאלות ברמת קושי בינונית וברמת מובחנות נמוכה, ובהתאם לתשובות הנבחן היא בוחרת עוד שאלות עד להתכנסות לכדי ציון סופי (למידע נוסף ראו מושינסקי ורונן, 1998). הבחינה מכוילת לציוני הבחינה הפסיכומטרית בנייר ובעיפרון. כאמור, מבחינת הנבחנים ללא לקות בחינת מפע״ם שקולה לבחינה פסיכומטרית רגילה (Moshinsky & Kazin, 2004). בחינת מפע״ם יועדה לנבחנים בתנאים מותאמים שנקבע כי הם זקוקים לתוספת זמן, ובבדיקה אמפירית נמצא שנבחנים בעלי לקויות קריאה ובעיות גופניות מסוימות משפרים בה את ציונם יותר מאשר נבחנים ללא לקויות (מושינסקי, 2002).

לבחינת מפע״ם כמה מאפיינים ייחודיים המתאימים להיבחנות בתנאים מותאמים:

א. הבחינה הממוחשבת בנויה כבחינת עוצמה, ובאופן תאורטי אין בה שום מגבלת זמן (רפ, רונן וכהן, 1996). בפועל, הקצבת הזמן לשאלה במפע״ם נקבעה על בסיס משכי פתרון השאלות שנמדדו בהעברות ניסיוניות של הבחינה. תוספת הזמן בבחינה נדיבה ביותר, והיא נעה בין 100% ל־400%. מאפיין זה של הבחינה מתאים במיוחד לנבחנים הזקוקים לתוספת זמן כפיצוי על לקות המאטה את תפקודם.

ב. כל שאלה מוצגת על מסך המחשב בנפרד ובהגדלה, והדבר מקל על הנבחן למקד את הקשב שלו בשאלה עצמה בלי שגירויים אחרים יסיחו את דעתו. נוסף על כך, קריאת מסכי ההסברים המופיעים בין סוגי השאלות ובין פרק לפרק אינה מוגבלת בזמן, מה שמאפשר הפסקות התרעננות קצרות בתוך הכיתה. מאפיינים אלו של הבחינה עשויים להתאים לנבחנים בעלי ADHD.

ג. סימון התשובות נעשה באמצעות המקלדת, ולכן הנבחן אינו נדרש להעביר את תשומת לבו מחוברת בחינה אל דף תשובות וחוזר חלילה.

בבחינת הניסוי במחקר זה הועבר נוסח של בחינת מפע״ם ששימש כבחינה אמתית במועדים קודמים.

שאלון משוב על ביצוע בחינת מפע״ם: נבחנים אשר נבחנו בבחינת המפע״ם מלאו שאלון משוב שנועד לבדוק את התוקף הנראה של בחינה זו. בשאלון נשאלו הנבחנים על תחושותיהם כלפי בבחינת המפע״ם (לדוגמה, עד כמה הם מרגישים שהישגם בבחינה הממוחשבת משקף את יכולתם). יש להדגיש כי משתתפי הניסוי מילאו את השאלון מיד לאחר ההיבחנות וקודם לקבלת הציון, וזאת כדי שהערכותיהם לא יושפעו מהציון שהשיגו במבחן.

בחינות הניסוי נערכו לאורך כשנה בארבעה מועדי בחינות רשמיים של הבחינה הפסיכומטרית.

זימון המשתתפים: הזמנה להשתתף בניסוי נשלחה לכל העומדים בקריטריונים מקרב הפונים למדור במו"ת ולכ־600 נבחנים ללא ADHD שנדגמו אקראית. בהזמנה הוצע לנבחנים להשתתף בבחינה פסיכומטרית ניסויית, שתיערך כשלושה שבועות לפני מועד הבחינה שאליה הם רשומים. כתמריץ להשתתפות הוסבר לנבחנים בהזמנה זו כי הבחינה תשמש להם תרגול לקראת הבחינה האמתית, ונוסף על כך הם יוכלו לאמוד בעזרת הציון שיקבלו בה את רמת הביצוע שלהם בבחינה הפסיכומטרית (הציון חושב ונמסר לנבחנים מיד בתום בחינת הניסוי).

לכל נבחן נשלחו שתי הזמנות זהות זו לזו, האחת בדואר והאחרת בדואר אלקטרוני. נוסף על כך, למועמדים בעלי ADHD פנה צוות המחקר גם בטלפון. מועמדים שהיו מעוניינים להשתתף בניסוי נרשמו באמצעות הדואר האלקטרוני או באמצעות הטלפון וקיבלו הזמנה אישית להשתתף במחקר. המשתתפים במחקר לא ידעו מראש לאיזו קבוצה שובצו, אך הובהר להם מבעוד מועד שהם עשויים להיבחן בבחינת ני"ע או בבחינה ממוחשבת.

נבחנים בעלי ADHD שנענו להזמנה להשתתף במחקר הוקצו אקראית לקבוצת הניסוי או לקבוצת ביקורת. משתתפים ללא ADHD הוקצו אקראית לאחת משתי קבוצות הביקורת האחרות. לנבחנים בעלי ADHD הומלץ ליטול תרופות ביום בחינת הניסוי אם תכננו לעשות כן ביום הבחינה האמתית, וזאת כדי למנוע מצב שבו הפער שבין ציוני בחינת המחקר והבחינה האמתית יושפע מנטילת התרופה. מסיבות לוגיסטיות הוקצו הנבחנים לקבוצות לפני הגעתם לאתר הבחינה. ביום הבחינה ביטלו חלק מהמוזמנים את השתתפותם, ולכן לא היו הקבוצות שוות בגודלן, דבר שלא היה קורה לו הוקצו הנבחנים לקבוצות בעת הגעתם.

לכל הנבחנים הובהר שהציון שיקבלו בבחינת הניסוי הוא רק אומדן של רמת ביצועיהם בבחינה האמתית, וכי להשתתפותם בניסוי לא תהיה כל השפעה לא על ציונם בבחינה האמתית ולא על התנאים שלהם יהיו זכאים בבחינה האמתית במקרה שהם בעלי ADHD.

חשוב לציין כי לעתים הבחינה בתנאים מותאמים נערכת כשבועיים לאחר הבחינה הפסיכומטרית הרגילה. לפיכך, בכמה מן המועדים נבחנו נבחנים בעלי ADHD וללא ADHD בבחינת הניסוי יחד, אך בבחינה האמתית הם נבחנו במועדים שונים זה מזה. במילים אחרות, פרק הזמן שבין בחינת המחקר לבחינה האמתית היה שונה מעט לגבי כל סוג נבחנים.

לא נאמר לנבחנים כי הבחינה הממוחשבת היא בחינת מפע"ם, והיא הוצגה להם כבחינה ממוחשבת ניסיונית. הדבר נעשה משום שרבים מהפונים בעלי ADHD פונים בבקשה להיבחן בבחינת מפע"ם, ויתכן כי לו היו יודעים בוודאות שבחינת הניסוי היא בחינת מפע"ם, היה הדבר משפיע על תפקודם בבחינה (למשל, משום שהמוטיבציה שלהם להצליח הייתה משתנה).

במהלך בחינת המחקר נשמרו תנאי היבחנות דומים ככל האפשר לתנאי הבחינה האמתית. נבחנים בעלי ADHD שזומנו לבחינת ני"ע מותאמת נבחנו באותם התנאים המותאמים שבהם נבחנו לאחר מכן בבחינה האמתית. התאמות אלו כללו הפסקות בין פרקי הבחינה וצמצום של מספר

הנבחנים בחדר (התאמה שאושרה ל־99% מהנבחנים). במקרה הצורך אושר לנבחנים גם להשתמש במחשבון המבצע את ארבע פעולות החשבון הבסיסיות (56%) ובגיליון תשובות מוגדל (58%). נבחנים בעלי ADHD שהיו רשאים להשתמש במחשבון בבחינה האמתית וזומנו לבחינת מפע״ם ניסויית היו רשאים להשתמש בו גם בבחינה זו.

מהלך הבחינה: לאחר הקראת הוראות הסבר כלליות על מבנה הבחינה קיבלו המשתתפים שנבחנו בבחינת ני״ע את חוברות הבחינה ואת הגיליונות לסימון התשובות, ואילו נבדקים שנבחנו בבחינת המפע״ם החלו את הבחינה על גבי צג המחשב. בסיום הבחינה התבקשו כל הנבחנים למלא שאלונים לאיתור ADHD, לשם קבלת מידע קליני נוסף. נבחנים בבחינת המפע״ם התבקשו למלא גם שאלון משוב. כמו כן נשאלו המשתתפים אם נטלו תרופה לטיפול בהפרעת קשב וריכוז במהלך הבחינה ואם הם מתכוונים ליטול תרופה כזו במהלך הבחינה האמתית. המענה על שאלות אלו – כמו על שאר השאלות בשאלון – לא היה בגדר חובה.

נבחני בחינת ני״ע החלו וסיימו את הבחינה יחדיו. במהלך זמן מילוי השאלון (כ־15 דקות) חושבו ציוניהם באמצעות שבלונות המונחות על גיליון התשובות ומאפשרות ספירת הסימונים הנכונים המהווים את ציוני הגלם ותוכנת מחשב לחישוב ציוני התקן. לאחר סיום מילוי השאלונים נמסרו הציונים לנבחנים. נבחנים בבחינת מפע״ם החלו גם הם את הבחינה יחדיו, אולם בשל מבנה הבחינה היה משכה שונה לכל נבחן. בתום הבחינה, אך קודם לקבלת הציון, התבקשו הנבחנים למלא את שאלוני הקשב והריכוז וכן את שאלון המשוב. בתום מילוי השאלונים קיבלו הנבחנים את ציונם. כל נבחן בכל אחת מהקבוצות קבל ארבעה ציונים: ציון כללי, ציון בתחום המילולי, ציון בתחום הכמותי וציון בתחום האנגלית. כל הנבחנים שהגיעו להיבחן בבחינת הניסוי סיימו את הבחינה.

מערך המחקר

מערך המחקר מוצג בלוח 1:

לוח 1 – מערך קבוצות המחקר והביקורת בניסוי

		ADHD	ללא ADHD
סוג בחינה	מפע"ם	קבוצה (1) – מחקר N=58	קבוצה (2) – ביקורת מפע"ם N=83
	ני"ע	קבוצה (3) – ביקורת ADHD (בחינת ני"ע מותאמת) N=61	קבוצה (4) – ביקורת רגילה (בחינת ני"ע רגילה) N = 56

מערך זה אפשר לבחון את השערת השיפור הדיפרנציאלי – כלומר, לבדוק אם ההיבחנות בבחינת מפע״ם תורמת לשיפור דיפרנציאלי של הישגי נבחנים בעלי ADHD, הן בהשוואה להישגיהם בבחינת ני״ע בתנאים מותאמים המועברת להם כיום והן בהשוואה לשיפור בהישגי נבחנים ללא ADHD. קבוצות הביקורת אפשרו גם לקזז מן ההערכה את השיפור בשל אפקט מבחן־מבחן חוזר ובשל השפעת המוטיבציה, וכך לבדוק את האפקט של בחינת המפע״ם בלבד.

(א) השערת השיפור הדיפרנציאלי: השוואת קבוצת המחקר לקבוצות הביקורת

ציוני קבוצת המחקר בבחינת הניסוי (נבחנים בעלי ADHD שנבחנו בבחינת המפע״ם) הושוו לציוני קבוצות הביקורת בעזרת ניתוח שונות רב־משתני כמפורט בהמשך. ציוני הבחינה האמתית שימשו כמשתנה מפקח (covariate) כדי לנטרל הבדלים בין רמות היכולת של הנבחנים. ניתוח זה נערך לשלושת ציוני התחומים – מילולי, כמותי ואנגלית – ולציון הכללי בבחינה. ניתוח זה מסתמך על ההנחה שציוני אוכלוסיות הנבחנים מתפלגים נורמלית ואין הבדל משמעותי בשונות הציונים בכל קבוצה (הנחה זו נבדקה לפני עריכת הניתוח). בהתאם לפרדיגמת השיפור הדיפרנציאלי נבדק אם יש אינטראקציה בין הקבוצה לסוג הבחינה (אוכלוסייה X מתן התאמות) על ההישגים בבחינה, וזאת בנטרול הבדלי היכולת בין הקבוצות אם הם קיימים. הבקרה על משתנה היכולת מאפשרת להסיק כי מקורם של הבדלים הוא באופי המבחן.

באופן דומה אפשר להשוות את ביצועי קבוצת הנבחנים הרגילים במפע״ם לביצועי קבוצת הנבחנים הרגילים בני״ע. בהתבסס על מחקרי העבר, לא ציפינו להבדל משמעותי בין הישגי קבוצות הנבחנים הרגילים. כל פער שיימצא בין קבוצות הנבחנים הרגילים אמור לשמש נקודת השוואה להבדלים בין קבוצות בעלי ADHD, אם יימצאו כאלה. לדוגמה, אם שתי קבוצות הנבחנים ללא ADHD יגיעו לאותם הישגים בבחינת ני״ע ובבחינת מפע״ם, ואילו באוכלוסיית בעלי ADHD יימצא יתרון להישגי הנבחנים במבחן מפע״ם, אפשר יהיה להסיק שבחינת המפע״ם משפרת דיפרנציאלית את ציונם של בעלי ADHD.

(ב) השוואת מדדי ביצוע בחינת מפע״ם בין נבחנים עם בעלי ADHD לנבחנים ללא ADHD

מדדי ההתכנסות של הנבחנים בעלי ADHD שנבחנו בבחינת המפע״ם (קבוצת הניסוי) הושוו למדדים של הנבחנים ללא ADHD שנבחנו בבחינת המפע״ם (קבוצת הביקורת). בפרט נבדקו מדדים אלו באמצעות בחינת קצב ההתכנסות וטיבה. קצב ההתכנסות נקבע לפי מספר התשובות לשאלות הנדרשות כדי להגיע לרמת ביטחון בציון, שנקבעה מראש בכל בחינה. טיב ההתכנסות נקבע באמצעות בדיקת המידה שבה משתנה טעות המדידה של אומדן היכולת. הניתוח כלל ניתוח שונות שבו טיב ההתכנסות היה המשתנה התלוי, וסוג הקבוצה (בעלי ADHD או ללא ADHD) שימש משתנה בין־נבדקי בלתי תלוי. הבדל משמעותי בין התוצאות של שתי הקבוצות שנבחנו במפע״ם יעיד על כך שבמפע״ם אופי הביצוע של נבחנים עם ADHD שונה מזה של נבחנים ללא ADHD. לבסוף נבדקו הבדלים בין הקבוצות בזמן התגובה לכל אחת מהשאלות.

(ג) שיפוטי מומחים וניתוח אידיוגרפי

שאלת המחקר הייתה אם שיפוטי מומחים עשויים לנבא ציונו של מי מן הנבחנים בעלי ADHD בבחינת מפע״ם יעלה על ציונו בבחינת נייר ועיפרון מותאמת, וציונו של מי מהם לא יעלה עליו או אף ייפול ממנו. לפיכך נעשה ניסיון לבדוק את תוקף שיפוטי המומחים (צוות היחידה לבחינות מותאמות במרכז הארצי לבחינות ולהערכה). לכל נבחן נבדקו התשובות לשתי שאלות: א. האם

לפי שיפוטי המומחים יעלה ציונו של הנבחן בבחינת מפע״ם על ציונו בבחינת נייר ועיפרון (להלן יפורטו הקריטריונים להחלטה זו)? ב. האם אמנם קרה הדבר בפועל?

כדי להשיב על השאלה הראשונה קראו שלוש מומחיות מצוות היחידה את כל תיקי הנבחנים שאובחנו כבעלי ADHD ונבחנו בניסוי בבחינת מפע״ם. לכל נבחן הזינו הקוראות נתונים בקידוד דיכוטומי (1=קיים, 2=לא קיים) על הממדים האלה: (א) מוסחות וחזרתיות; (ב) בעיות בארגון ובתכנון; (ג) איטיות כללית בתפקוד. לפי שקלול התפקוד בכל הממדים האלה קבעה כל קוראת אם לדעתה ציונו של הנבחן בבחינת מפע״ם יהיה גבוה מזה שיקבל בבחינת ני״ע מותאמת, וזאת על בסיס הערכתה אם הארגון המובנה של הבחינה הממוחשבת או מסגרת הזמן המרווחת שלה עתידים לסייע לו. נקבע שציונו של נבחן בבחינה ממוחשבת עשוי לעלות על ציונו בבחינת ני״ע אם שתי קוראות או יותר קבעו זאת באופן חד משמעי. בדומה לכך נקבע שציונו של נבחן בבחינה ממוחשבת אינו עשוי לעלות על ציונו בבחינת ני״ע אם שתי קוראות או יותר קבעו זאת. קריאת התיקים נעשתה לאחר הבחינות, אך הקוראות לא נחשפו לנתוני הבחינות של הנבחן.

כדי להשיב על השאלה השנייה חושב הפער בין תפקודו של כל נבחן בקבוצת הניסוי (קבוצה 1) בבחינת הניסוי האדפטיבית לתפקודו בבחינה האמתית (בגרסת נייר ועיפרון מותאמת). פער חיובי ביטא תפקוד טוב יותר בבחינת המפע״ם הניסויית מאשר בבחינה האמיתית. חשוב לציין כי ללא קשר לסוג הבחינה יש לצפות לציון גבוה יותר בבחינה האמתית, וזאת מהסיבות שפורטו לעיל, ולכן פער שלילי או קרוב לאפס אין פירושו בהכרח שהנבחן לא הסתייע בבחינת המפע״ם. עם זאת, בחרנו לראות רק בפערים חיוביים עדות לשיפור משמעותי בבחינה זו. בדיקת התפלגות הפערים אפשרה לנו לבחון את תפקודו של כל נבחן ונבחן ולבדוק תפקודם של אילו נבחנים השתפר בבחינת המפע״ם ותפקודם של אילו נבחנים הורע או לא השתנה בה. מבחני t נערכו כדי לבדוק את מובהקות ההבדלים הממוצעים בין הנבחנים שהקוראות קבעו שציונם עשוי להשתפר בבחינת מפע״ם לבין הנבחנים שהקוראות קבעו שציונם לא ישתפר בבחינת מפע״ם.

(ד) נתונים תיאוריים מתוך התשובות לשאלון המשוב לגבי בחינת מפע״ם

חושבו שכיחויות, ממוצעים וסטיות תקן של המשתנים שהופיעו בשאלון המשוב לגבי בחינת מפע״ם. כמו כן נבדק הקשר בין הפער בהישגים בפועל בין בחינת ני״ע ובחינת מפע״ם לבין הדיווח העצמי על שביעות הרצון מהבחינה, מידת העדפת הבחינה הממוחשבת על פני בחינת נייר ועיפרון ומידת ההצלחה הנתפסת בבחינת מפע״ם לעומת בחינת ני״ע.

ממצאים

נתונים תיאוריים

בלוח 2 מוצגים הממוצעים וסטיות התקן של ציוני הנבחנים בבחינת מפע״ם ובבחינת נייר ועיפרון של נבחנים בעלי ADHD וללא ADHD. כמו כן מוצגים נתוני הבחינה האמתית של שתי הקבוצות.

לוח 2 – ציוני הנבדקים משתי קבוצות הניסוי בבחינת הניסוי ובבחינה האמתית

קבוצה	בחינת מפע"ם				בחינת ני"ע			
	ADHD (N=58)		No ADHD (N=83)		ADHD (N=61)		No ADHD (N=56)	
	בחינת ניסוי							
	ממוצע	ס"ת	ממוצע	ס"ת	ממוצע	ס"ת	ממוצע	ס"ת
חשיבה מילולית	110	20	116	16	108	16	114	15
חשיבה כמותית	111	18	119	17	108	19	115	18
אנגלית	113	24	115	23	113	21	116	21
ציון כללי	**561**	**101**	**597**	**81**	**551**	**88**	**585**	**86**
	בחינה אמתית							
	ממוצע	ס"ת	ממוצע	ס"ת	ממוצע	ס"ת	ממוצע	ס"ת
חשיבה מילולית	108	20	117	18	110	16	114	16
חשיבה כמותית	109	17	117	18	112	17	117	16
אנגלית	117	23	117	22	116	22	117	22
ציון כללי	**558**	**100**	**595**	**87**	**569**	**83**	**589**	**86**

מן הלוח עולה כי יש הבדלים בין ציוני הנבחנים בקבוצות המחקר השונות, הן בבחינת מפע״ם והן בבחינת הנייר והעיפרון: בשני סוגי הבחינות ציונם של נבחנים ללא ADHD גבוה מזה של בעלי ADHD. משמעותם הסטטיסטית של הבדלים אלו תידון בהמשך.

כמו כן, בכמה מהמקרים יש הבדלים בין ציוני בחינת הניסוי לציוני הבחינה האמתית. כיוונם של הבדלים אלו אינו עקבי: בכמה מהמקרים השינוי בציון חיובי, ובכמה מהם – שלילי. לשם בדיקת הקשר בין ציוני בחינת הניסוי לציוני הבחינה האמתית חושבו מתאמי פירסון בין ציוני הבחינות בכל אחת מקבוצות המחקר (בעלי ADHD וללא ADHD). מתאמים אלו מוצגים בלוח 3.

לוח 3 – מתאמים בין הציונים של נבדקי הניסוי בבחינת הניסוי ובבחינה האמתית

קבוצה	בחינת מפע"ם		בחינת ני"ע		כלל הנבחנים (N=258)
	ADHD (N=58)	No ADHD (N=83)	ADHD (N=61)	No ADHD (N=56)	
חשיבה מילולית	0.88	0.85	0.79	0.81	0.84
חשיבה כמותית	0.85	0.83	0.76	0.82	0.81
אנגלית	0.90	0.92	0.93	0.92	0.91
ציון כללי	**0.94**	**0.92**	**0.88**	**0.93**	**0.92**

מלוח 3 עולה כי המתאמים שבין ציוני בחינת הניסוי לציוני הבחינה האמתית גבוהים הן לגבי נבחנים שנבחנו בניסוי בבחינת מפע״ם והן לגבי נבחנים שנבחנו בניסוי בבחינת נייר ועיפרון (המתאם בין ציוני היבחנות ראשונה לשנייה בקרב כלל האוכלוסייה הוא 0.89). עם זאת יש לציין כי מרווח הזמן בין ההיבחנויות היה כ־3–5 שבועות, ואילו נתונים לגבי היבחנות חוזרת באוכלוסייה הכללית מחושבים על פער של חצי שנה בקירוב. נוסף על כך, בקרב קבוצת הנבחנים בעלי ADHD היה המתאם בין הבחינות כשבחינת הניסוי הייתה בחינת מפע״ם גבוה מן המתאם ביניהן כשבחינת הניסוי הייתה בחינת נייר ועיפרון, וזאת בתחומים המילולי והכמותי (בציון המילולי 0.88 לעומת 0.79, $Z=1.62$, $p=0.05$, ובציון הכמותי 0.85 לעומת 0.76, $Z=1.38$, $p=0.08$, בהתאמה) ובציון הכללי (0.94 לעומת 0.88, $Z=1.92$, $p=0.05$). זאת אף על פי שבחינת הנייר והעיפרון נערכה בתנאים זהים לתנאי הבחינה האמתית, בעוד שבחינת המפע״ם הייתה שונה ממנה.

הבדלים בין קבוצת המחקר (בעלי ADHD) לקבוצת הביקורת (ללא ADHD) בציוני בחינת מפע"ם לעומת בחינת נייר ועיפרון

כדי לבחון את מובהקותם הסטטיסטית של ההבדלים בין ציוני קבוצות המחקר והביקורת בבחינת מפע״ם ובין ציוניהן בבחינת נייר ועיפרון תוקנו הציונים בעזרת ניכוי יכולתם האמתית של הנבחנים, וזאת בהתבסס על ציוני הבחינה האמתית. ציוני בחינת הניסוי המתוקנים של שתי הקבוצות מוצגים בלוח 4.

לוח 4 – ציונים מתוקנים בבחינת הניסוי

קבוצה	בחינת מפע"ם				בחינת ני"ע			
	ADHD (N=58)		No ADHD (N=83)		ADHD (N=61)		No ADHD (N=56)	
	ממוצע מתוקן	טעות תקן	ממוצע מתוקן	טעות תקן	ממוצע מתוקן	טעות תקן	ממוצע מתוקן	טעות תקן
חשיבה מילולית	113	1.2	113	1.0	110	1.2	112	1.2
חשיבה כמותית	115	1.4	116	1.2	109	1.4	113	1.4
אנגלית	113	1.2	115	1.0	113	1.2	116	1.2
ציון כללי	**581**	**4.6**	**583**	**3.8**	**560**	**4.5**	**576**	**4.7**

לשם בדיקת מובהקות ההבדלים נערכו ניתוחי שונות דו־גורמיים (ANCOVA). סוג הקבוצה וסוג הבחינה שימשו כמשתנים הבלתי תלויים, וציוני בחינת הניסוי בתחומים המילולי, הכמותי והאנגלית, וכן והציון הכללי בה, היו המשתנים התלויים. ממצאי ניתוחים אלו מסוכמים בלוח 5.

לוח 5 – ניתוח שונות לבדיקת השפעת סוג הבחינה והקבוצה על הציונים בתחומי הבחינה ועל הציון הכללי בבחינת הניסוי

ערך F				
מקור / תחום	מילולי	כמותי	אנגלית	כללי
קבוצה	0.6	4.0*	5.2*	4.0*
סוג בחינה	2.7	11.4*	0.5	9.7*
קבוצה X סוג בחינה	2.1	0.6	0.0	2.5

* הבדל מובהק ברמת מובהקות 0.05

כפי שעולה מלוח 5, לגבי הציון בתחום המילולי לא נמצאו הבדלים מובהקים סטטיסטית בין רמות המשתנים קבוצה, סוג בחינה, ואינטראקציה בין קבוצה לסוג בחינה. בהתאם להשערת השיפור הדיפרנציאלי נמצא כי בעוד שאין פער בין ציוני קבוצת הביקורת לקבוצת המחקר בבחינת מפע״ם (113), יש פער קטן בין ציוני שתי הקבוצות בבחינת הנייר והעיפרון, ונבחנים בעלי ADHD קיבלו בה ציון נמוך יותר מנבחני הביקורת (110 לעומת 112). ואולם, הבדל זה אינו מובהק סטטיסטית.

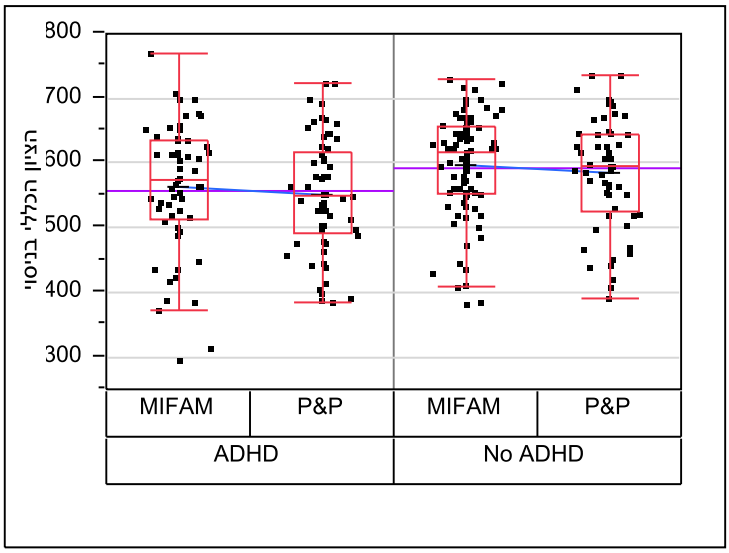
לגבי הציון בתחום הכמותי נמצאו הבדלים מובהקים סטטיסטית בין רמות המשתנים קבוצה וסוג בחינה: בתחום בחינה זה קיבלו נבחנים ללא ADHD ציון גבוה יותר מנבחנים בעלי ADHD. נוסף על כך, בבדיקה מעבר לקבוצות היה ציונם של נבחנים בבחינת מפע״ם בתחום הכמותי גבוה מציונם של נבחנים בבחינת נייר ועיפרון. עם זאת, האינטראקציה בין הקבוצה לסוג הבחינה לא הייתה מובהקת סטטיסטית. בהתאם להשערת השיפור הדיפרנציאלי הפער בין ציוני קבוצת הביקורת לציוני קבוצת המחקר בבחינת מפע״ם היה קטן מהפער בין ציוני שתי הקבוצות בבחינת הנייר והעיפרון (115 לעומת 116, מול 109 לעומת 113), ונבחנים בעלי ADHD קיבלו בה ציון נמוך יותר מנבחני הביקורת, אולם גם הבדל זה אינו מובהק סטטיסטית.

לגבי הציון בתחום האנגלית נמצאו הבדלים מובהקים בין רמות משתנה הקבוצה בלבד: בתחום בחינה זה קיבלו נבחנים ללא ADHD ציון גבוה יותר מנבחנים בעלי ADHD. שאר ההבדלים לא נמצאו מובהקים סטטיסטית.

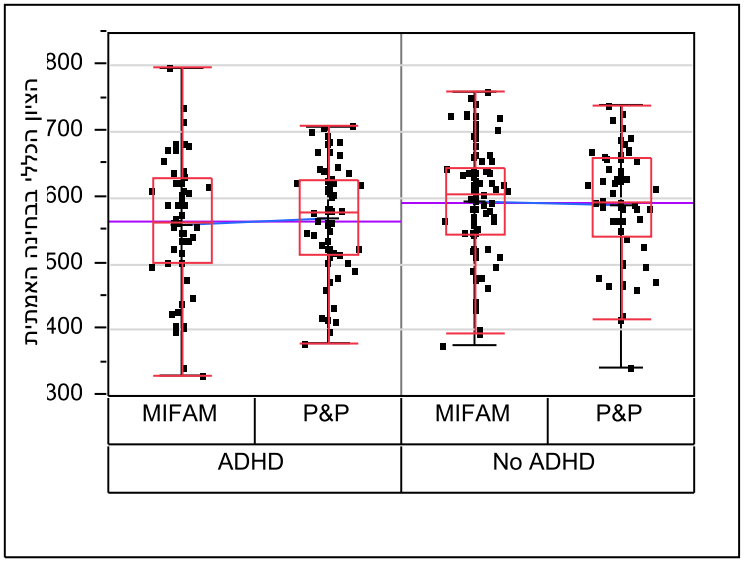
לבסוף, לגבי הציון הכללי בבחינה נמצאו הבדלים מובהקים סטטיסטית בין רמות המשתנים קבוצה וסוג בחינה: נבחנים ללא ADHD קיבלו ציון כללי גבוה יותר מנבחנים בעלי ADHD. נוסף על כך, בבדיקה מעבר לקבוצות היה ציונם הכללי של נבחנים בבחינת מפע״ם גבוה משל נבחנים בבחינת נייר ועיפרון. האינטראקציה בין הקבוצה לסוג הבחינה לא הייתה מובהקת סטטיסטית. בהתאם להשערת השיפור הדיפרנציאלי, הפער בין ציוני קבוצת הביקורת לקבוצת המחקר

בבחינת מפע״ם (2 נקודות) קטן בהרבה מהפער בין ציוני שתי הקבוצות בבחינת הנייר ועיפרון (16 נקודות), בחינה שבה קיבלו נבחנים בעלי ADHD ציון נמוך יותר מנבחני הביקורת. עם זאת, גם לגבי הציון הכללי הבדל זה אינו מובהק סטטיסטית. תוצאות ניתוחים אלו מוצגות באיורים 1 ו־2.

איור 1 – ציוני קבוצות הניסוי והמחקר בבחינת הניסוי (מפע"ם מול נייר ועיפרון)



איור 2 – ציוני קבוצות הניסוי והמחקר בבחינה האמתית על פי סיווגן בניסוי



לסיכום, מכלול הנתונים אינו תומך בהשערת השיפור הדיפרנציאלי, והוא מראה כי לא נמצאה אינטראקציה מובהקת סטטיסטית בין סוג הבחינה לקבוצת המחקר. עם זאת ראוי לציין כי כיוון הבדלי הממוצעים תומך בהשערת האינטראקציה, וכי השונות הגדולה בקרב נבחנים בעלי ADHD, בעיקר בתפקודם בבחינתם מפע״ם, וכן גודל המדגם הקטן, היו עלולים לפגוע בעוצמת הבדיקה ולהסתיר הבדל קטן. סוגיה זו תידון ביתר הרחבה בהמשך.

כדי לבחון את ההבדלים בין התנהלותם של נבחנים בעלי ADHD לנבחנים ללא ADHD בבחינת מפע״ם (ניסוי בלבד) הושוו כמה פרמטרים המופקים ממדדי הבחינה של שתי הקבוצות. בפרט, בבחינת מפע״ם אפשר לאמוד את טעות המדידה של יכולת הנבחן (PVAR), לחשב את אחוז ההתכנסות במבחנים וכן להשוות בין זמני התגובה לכל שאלה. טעות המדידה המגולמת באומדן של כל רמת יכולת של נבחנים נמצאת ביחס הפוך לסכום האינפורמציה של השאלות שנשאלו ברמה זו. האינפורמציה של כל שאלה היא ערך המייצג את תרומתה להערכת היכולת של הנבחנים. סף ההתכנסות בכל תחום נגזר מהמהימנות שהתקבלה בכל תחום בהעברה בנייר ובעיפרון. אחוז ההתכנסות הוא אחוז המבחנים שהתכנסו בכל קבוצה בניסוי (למידע נוסף, ראה Baker & Frank, 2001).

שיערנו כי טעות המדידה תהיה גבוהה בקרב נבחנים בעלי ADHD יותר מבקרב הנבחנים האחרים, וזאת בשל תנודתיות בתפקודם בבחינה. שיערנו גם כי אחוז המתכנסים בקרב נבחנים בעלי ADHD יהיה נמוך יותר מבקרב הנבחנים האחרים, וזאת בשל חוסר עקביות בתשובותיהם על שאלות קלות וקשות. שיערנו כי בשל קושי זה של המבחן להגיע להתכנסות בקרב נבחנים בעלי ADHD יהיה מספר השאלות שיוצגו להם בבחינה גדול יותר. לבסוף, שיערנו כי זמני התגובה של נבחנים בעלי ADHD יהיו ארוכים יותר משל הנבחנים האחרים. להלן יוצגו ממצאים מפורטים בכל אחד ממדדים אלו.

טעות המדידה המגולמת באומדן היכולת: כדי לבחון את ההשערה כי טעות המדידה המגולמת באומדן היכולת גבוהה יותר בקרב נבחנים עם ADHD, חושבו ציונים ממוצעים של טעות המדידה לשתי הקבוצות. התפלגות טעות המדידה אינה נורמלית ואינה לוג־נורמלית, ולכן, לצורך בדיקת מובהקותם הסטטיסטית של ההבדלים בין הקבוצות, נעשה שימוש במבחן אי־פרמטרי (Wilcoxon Rank Sums) לשם בדיקת מדד המרכז, ובמבחן Levene – לשם בדיקת שוויון השונויות (השערה חד־צדדית הן למדד המרכז והן למדד הפיזור הבודקת אם טעות המדידה אצל נבחנים בעלי ADHD שווה לטעות המדידה אצל נבחנים ללא ADHD או נמוכה ממנה). תוצאות השוואה זו מוצגות בלוח 6, ומהן עולה כי בניגוד להשערה, אין הבדלים מובהקים בטעות המדידה בין נבחנים בעלי ADHD לבין נבחנים ללא ADHD.

לוח 6 – טעות המדידה המגולמת באומדן היכולת (PVAR*) בבחינת מפע"ם בשתי הקבוצות

קבוצה	בחינת מפע"ם				Wilcoxon	Levene
	ADHD (N=58)		No ADHD (N=83)		Prob > F	p-Value
	ממוצע	ס"ת	ממוצע	ס"ת		
חשיבה מילולית	0.1516	0.0348	0.1447	0.0222	0.1152	0.0354
חשיבה כמותית	0.1053	0.0273	0.1017	0.0268	0.1048	0.1632
אנגלית	0.0752	0.0168	0.0834	0.0354	0.7696	0.9930

*סף ההתכנסות: בתחום המילולי: pvar<=0.14, בתחום הכמותי: pvar<=0.1, באנגלית: pvar<=0.08

מספר המתכנסים ואחוז ההתכנסות: כדי לבחון את ההשערה כי בקרב נבחנים בעלי ADHD מספר המתכנסים ואחוז המתכנסים נמוכים יותר, חושבו אחוזי ההתכנסות שתי הקבוצות. לצורך בדיקת מובהקותם הסטטיסטית של ההבדלים בין הקבוצות נערך מבחן Adjusted Wald Test (השערה חד־צדדית), המתאים לבדיקת השערות על פרופורציות של שני מדגמים בלתי תלויים. תוצאות השוואות אלו מוצגות בלוח 7.

לוח 7 – מספר המתכנסים ואחוז ההתכנסות בבחינת מפע"ם בשתי הקבוצות*

	No ADHD (N=83)		ADHD (N=58)		
Prob.	%	N	%	N	
0.2188	65%	54	59%	34	חשיבה מילולית
0.2960	80%	66	76%	44	חשיבה כמותית
0.8170	88%	73	93%	54	אנגלית

*סף ההתכנסות: בתחום המילולי: pvar=0.14, בתחום הכמותי: pvar=0.1 , באנגלית: pvar=0.08

מהשוואת השכיחויות המוצגות בלוח עולה כי בפרקי חשיבה מילולית וחשיבה כמותית אחוז המתכנסים נמוך מעט יותר אצל נבחנים בעלי ADHD, אך הבדל זה אינו מובהק סטטיסטית. בפרק האנגלית אחוז המתכנסים בקבוצת ה־ADHD גבוה יותר.

כדי לבחון את ההשערה שנבחנים בעלי ADHD מקבלים מספר רב יותר של שאלות הושוו מספרי השאלות שקיבלו הנבחנים בכל קבוצה ובכל פרק. התפלגות מספר השאלות אינה נורמלית, ולכן לצורך זה נעשה שימוש במבחן Wilcoxon (Rank Sums) חד־צדדי (2-sample test, normal approximation). הממצאים מוצגים בלוח 8.

לוח 8 – מספר השאלות שקיבלו הנבחנים בבחינת מפע"ם בשתי הקבוצות

	No ADHD (N=83)		ADHD (N=58)		
Prob > Z	ס"ת	ממוצע	ס"ת	ממוצע	
0.1852	2.3	34.4	2.3	34.8	חשיבה מילולית
0.0148*	2.0	29.7	2.0	30.4	חשיבה כמותית
0.6613	2.6	23.0	2.5	22.9	אנגלית

גם כאן, בניגוד להשערה, בפרקי החשיבה המילולית ובפרקי האנגלית לא נמצאו הבדלים מובהקים סטטיסטית בין מספר השאלות שקיבלו נבחנים בעלי ADHD למספר השאלות שקיבלו נבחנים ללא ADHD. בפרקי החשיבה הכמותית נמצא כי נבחנים בעלי ADHD קיבלו שאלות רבות יותר, וזאת באופן מובהק סטטיסטית.

זמני תגובה: כדי לבחון את ההבדל בזמני התגובה בין קבוצת הניסוי לקבוצת הביקורת נערך לכל שאלה מבחן סטטיסטי התואם את סוג ההתפלגות שלה. נמצא כי במובהק, בכל התחומים ובכל סוגי השאלות ממוצע זמן התגובה על שאלה בקבוצת הנבחנים בעלי ADHD גבוה יותר, למעט

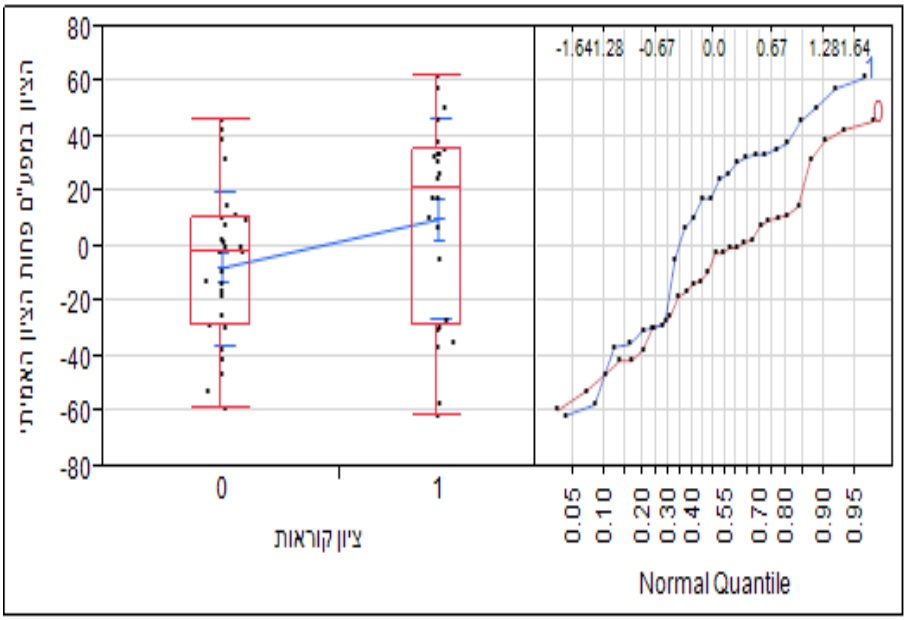
הזמן שהוקדש לקריאת קטע הקריאה באנגלית. הנבחנים בעלי ADHD השתמשו בממוצע ב־ 107%–145% מהזמן שבו השתמשו נבחני הביקורת למענה על שאלה.

שיפוטי מומחים לגבי השיפור הצפוי בבחינת מפע"ם

כדי לבחון אם צוות היחידה לבחינות מותאמות יכול להבחין הבחנה ראשונית בין נבחנים בעלי ADHD אשר ציונם בבחינת מפע״ם צפוי להיות גבוה מציונם בבחינתם האמתית לבין נבחנים בעלי ADHD שלגביהם אין הדבר כך, נערכה הבדיקה שתוארה בפרק שיטת המחקר.

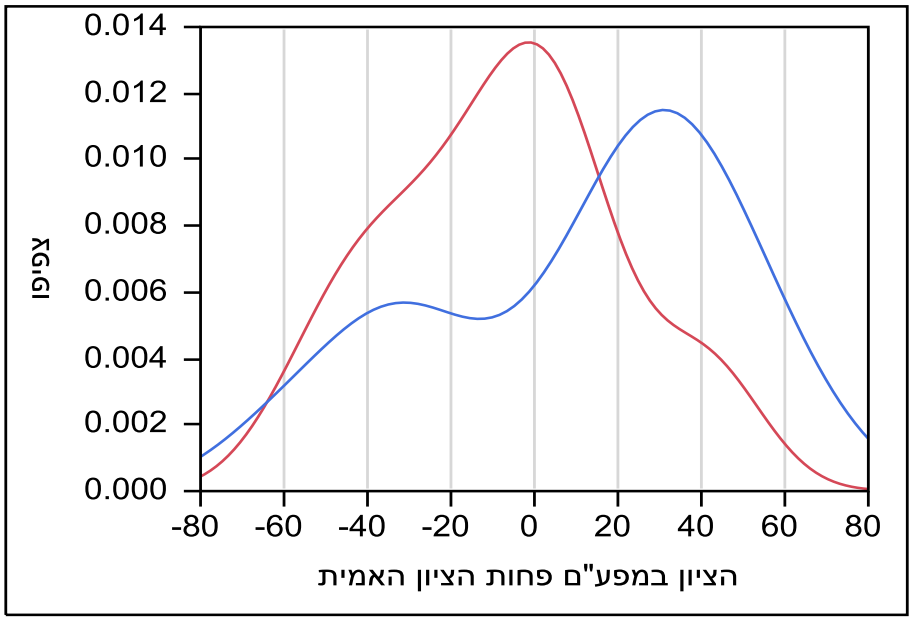
הממצאים הראו כי שיפוטי המומחים הצליחו להבחין בין נבחנים שציונם בבחינת מפע״ם היה גבוה מציונם בבחינת נייר ועיפרון (M=10.42, SD=36.4) לבין נבחנים שלגביהם לא היה הדבר כך (M=-7.21, SD=28). מבחן t חד זנבי העלה תוצאה מובהקת והראה כי ההפרש בין ציוני הבחינות בפועל (p=0.03) אצל נבחנים שהצוות קבע שציונם בבחינת מפע״ם יהיה גבוה יותר אמנם גדול יותר מאצל הנבחנים שהצוות קבע שציונם לא יהיה גבוה יותר. התפלגויות הפרשי הציונים של הנבחנים בניתוח זה כפונקציה של שיפוט המומחים מופיעות באיור 3.

איור 3 – הבדלים בין נבחנים שנקבע שציונם במפע"ם יהיה גבוה יותר לבין נבחנים שנקבע שציונם לא יהיה גבוה יותר



באיור 4 מוצגות צפיפויות ההתפלגות של שתי הקבוצות.

איור 4 – התפלגות ציוני הבחינה לפי חלוקת שיפוטי המומחים



באיור 3 משתקף ההבדל בין ממוצעי הפרשי הציונים של נבחנים שלגביהם נקבע כי ציונם בבחינת מפע״ם יהיה גבוה יותר ובין ממוצעי הפרשי הציונים של נבחנים שלגביהם נקבע כי ציונם בבחינת מפע״ם לא יהיה גבוה יותר. נוסף על כך האיור מציג את ההבדלים בהתפלגויות שתי הקבוצות. בקבוצת הנבחנים שהמומחים שיערו שתפיק תועלת מבחינת מפע״ם יש תת־קבוצה מובחנת שציונה בבחינה זו אמנם היה גבוה יותר מציונם בבחינת נייר ועיפרון, ותת־קבוצה מובחנת אחרת שציונה בבחינת מפע״ם היה נמוך במידה ניכרת. לעומת זאת בקבוצה השנייה התפלגות הציונים רציפה יותר, אם כי גם בה יש נבחנים שהפיקו תועלת ממפע״ם ונבחנים שציונם הורע בעקבותיה. קבוצה ייחודית זו של נבחנים – שנקבע שציונם ישתפר אך בפועל הורע ציונם – היא קבוצה הראויה למחקר מעמיק, וזאת לשם איתור המשתנים שהובילו לפער זה ולשם חיפוש אחר מאפיינים המשותפים לחבריה. מאיור 4 עולה כי אמנם יש חפיפה כלשהי בין ההתפלגויות, אך אפשר לזהות שתי התפלגויות מובחנות הנבדלות זו מזו בממוצעיהן.

תוצאות משוב מפע"ם

כדי לבדוק את שביעות רצונם של הנבחנים בבחינת מפע״ם, את ההבדלים בחוויית ההיבחנות בינה לבין בחינת נייר ועיפרון ואת ההבדלים אפשריים שבין הקבוצות נותחו נתוני שאלון המשוב. הניתוח נגע לנוחות השימוש במחשב, למידת המוטיבציה בבחינה, למידת שביעות הרצון מהמאפיינים הייחודיים של בחינת מפע״ם ולהבדלים בינה לבין בחינת נייר ועיפרון. נתונים אלו מוצגים בלוחות שלהלן.

נגישות הבחינה הממוחשבת

בלוח 9 מוצגים נתונים לגבי תדירות השימוש במחשב של הנבחנים ולגבי בהירות ההסברים בבחינה, וזאת בחלוקה לקבוצת המחקר ולקבוצת הביקורת.

לוח 9 – נתונים לגבי תדירות השימוש במחשב ומידת בהירות ההסברים בבחינה הממוחשבת לפי קבוצה

מספר פעמים בשבוע	ADHD (n=58)	ללא ADHD (n=83)	סה"כ (n=141)
תדירות השימוש במחשב בשבוע שלפני הבחינה			
3-6+	79%	84%	83%
0-2	21%	16%	17%
מידת הבהירות של ההסברים בבחינה הממוחשבת			
די ברורים – ברורים מאוד	85%	100%	98%
לא כל כך ברורים – מאוד לא ברורים	5%	0%	2%

כפי שעולה מן הלוח, רוב הנבחנים משתמשים במחשב בתדירות גבוהה ללא קשר לקבוצה שהם משתייכים אליה. רק מיעוטם הקטן (7%) לא השתמש במחשב כלל בשבוע האחרון. לא נמצאו הבדלים בין הקבוצות במידת השימוש במחשב ($\chi^2=0.94,\ p>0.05$). כ־98% מהנבחנים ציינו כי ההסברים היו די ברורים או ברורים מאוד, ובעיני נבחנים בעלי ADHD הייתה מידת הבהירות של ההסברים ($\chi^2=8.68,\ p<0.05$) נמוכה במובהק מבעיני נבחנים ללא ADHD.

מידת המאמץ שהושקע בביצוע בחינת הניסוי

הנבחנים התבקשו לדרג את מידת המאמץ שהשקיעו בבחינת הניסוי הממוחשבת בסולם בן 5 דרגות, מ״ללא מאמץ כלל״ עד ״מידה רבה מאוד״. התפלגות התשובות של כל קבוצה מוצגת בלוח 10.

לוח 10 – מידת המאמץ שהשקיעו הנבחנים בבחינת הניסוי הממוחשבת לפי קבוצה

	ADHD (n=58)	ללא ADHD (n=83)	סה"כ (n=141)
מידת המאמץ שהושקע בבחינת הניסוי הממוחשבת			
מידה רבה – מידה רבה מאוד	60%	68%	65%
מידה בינונית	35%	21%	27%
מידה מועטה – ללא מאמץ כלל	5%	11%	8%

מן הלוח עולה שרוב הנבחנים דיווחו שהשקיעו בבחינת הניסוי מאמץ במידה בינונית ומעלה ואף במידה רבה ומעלה. מכך אפשר להסיק כי רמת המוטיבציה של מרבית הנבחנים הייתה גבוהה, וכי ציוני הבחינה משקפים את יכולתם היטב. לא נמצאו הבדלים בין הקבוצות בשכיחות התשובות (χ^2=4.01, p>0.05).

תחושת הקושי בבחינה

בשאלון המשוב התבקשו הנבחנים להעריך את מידת הקושי הנתפסת של השאלות בבחינת מפע"ם בתחום המילולי, בתחום הכמותי ובתחום האנגלית. בחינת מפע"ם מציגה לנבחן שאלות בהתאם לרמת היכולת שהפגין כשענה על השאלות הקודמות, ולכן הוא נתקל בשאלות רבות המותאמות ליכולתו ובשאלות מעטות בלבד שהן קלות מאוד עבורו, וזאת בניגוד לבחינת נייר ועיפרון, שבה הנבחן מקבל שאלות בכל רמות הקושי. מסיבה זו רמת הקושי הנתפסת של השאלות בבחינת מפע"ם עשויה להיות גבוהה יחסית. התפלגות רמת הקושי הנתפסת של השאלות (מ"קלות מאד" ועד "קשות מאד") מוצגת בלוח 11, הן לנבחנים בעלי ADHD והן לנבחנים ללא ADHD. שיעורי הנבחנים בסך הכול, מעבר לקבוצות, מוצגים מתחת לשיעוריהם בכל תחום.

לוח 11 – מידת הקושי הנתפסת של שאלות הבחינה לפי תחומים ולפי קבוצות

	קלות מאוד		די קלות		בינוניות		די קשות		קשות מאוד	
	ADHD	ללא ADHD	ADHD	ללא ADHD	ADHD	ללא ADHD	ADHD	ללא ADHD	ADHD	ללא ADHD
חשיבה מילולית (%)	0%	1%	10%	13%	52%	52%	33%	30%	5%	4%
סה"כ (n=141)	**0.5%**		**12%**		**52%**		**32%**		**4%**	
חשיבה כמותית (%)	2%	0%	10%	18%	38%	47%	38%	28%	12%	7%
סה"כ (n=141)	**0.5%**		**15%**		**43%**		**32%**		**9%**	
אנגלית (%)	3.5%	1%	10%	13%	41%	39%	31%	34%	14%	13%
סה"כ (n=141)	**2%**		**12%**		**40%**		**33%**		**13%**	

מן הלוח עולה שרוב הנבחנים העריכו שמידת הקושי של השאלות בתחום המילולי בינונית (52%), ושמיעוטן הן קשות מאוד או קלות מאוד. לא נמצאו הבדלים בין נבחנים בעלי ADHD לנבחנים רגילים בכל הנוגע למידת הקושי שהם מייחסים לשאלות בתחום המילולי (χ^2=1.57, p>0.05).

בתחום הכמותי העריכו רוב הנבחנים שהשאלות בינוניות, די קלות ודי קשות (43%, 15% ו־32% בהתאמה), ושמיעוטן הן קשות מאוד או קלות מאוד. שיעור הנבחנים בעלי ADHD שהעריכו ששאלות בתחום זה מצויות בתחום הקשה (קשות, קשות מאוד) עולה על שיעור הנבחנים ללא ADHD שהעריכו הערכה כזו, אולם הבדל זה אינו מובהק סטטיסטית (χ^2=5.82, p>0.05).

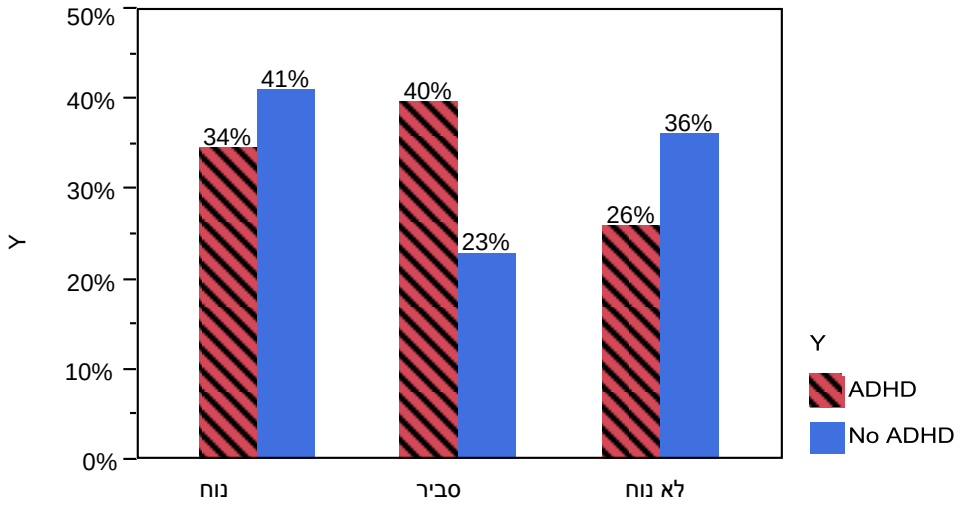
באנגלית העריכו רוב הנבחנים שהשאלות ברמת קושי בינונית ודי קשה (40% ו־33% בהתאמה). כמה מהם העריכו שהשאלות קשות מאוד (13%) ומיעוטם העריך שהן קלות מאוד. גם בתחום האנגלית אין הבדלים בין נבחנים בעלי ADHD לנבחנים ללא ADHD בכל הנוגע למידת הקושי שהם מייחסים לשאלות (χ^2=1.19, p>0.05).

שביעות רצון ממאפייני המבחן

הנבחנים התבקשו לחוות את דעתם על השפעת מאפיינים ייחודיים של בחינת מפע״ם על רמת הנוחות של הנבחן. הנבחנים ציינו עד כמה נוחים או לא נוחים היו בעיניהם הצפייה הממושכת במסך, ההצגה של כל שאלה במסך נפרד, הקצבת הזמן הנפרדת לכל שאלה והעובדה שאי אפשר לחזור לשאלות קודמות. התפלגות התשובות מוצגת באיורים 5–8, ובעזרתם אפשר להשוות בין נבחנים בעלי ADHD לנבחנים ללא ADHD.

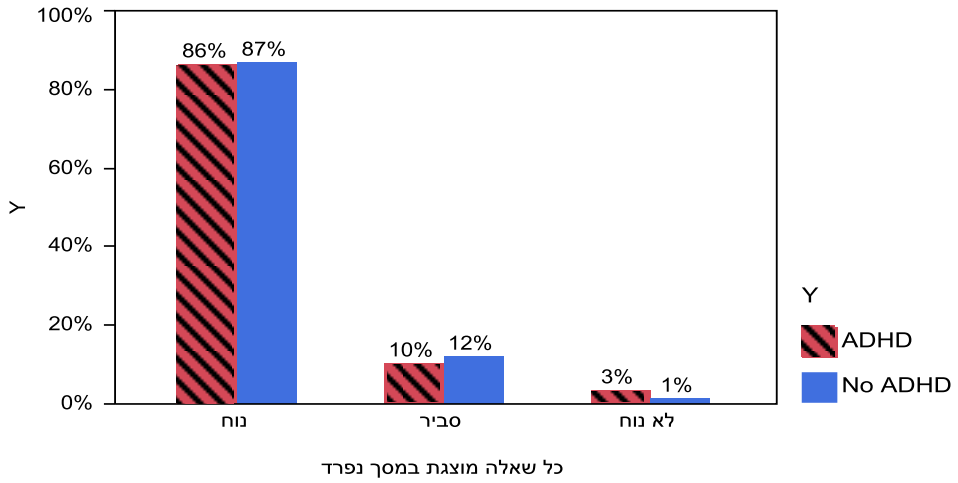
כשני שלישים מן הנבחנים (66%) חשו כי הצפייה הממושכת במסך המחשב היא סבירה עד נוחה. אחוז קטן יותר מהנבחנים בעלי ADHD חש כי הצפייה הממושכת במסך אינה נוחה לו, אך הבדל זה לא היה מובהק סטטיסטית (χ^2=4.69, p>0.05). ממצאים אלו מוצגים באיור 5.

איור 5 – התפלגות שביעות הרצון מן הצפייה הממושכת במסך בבחינת מפע"ם



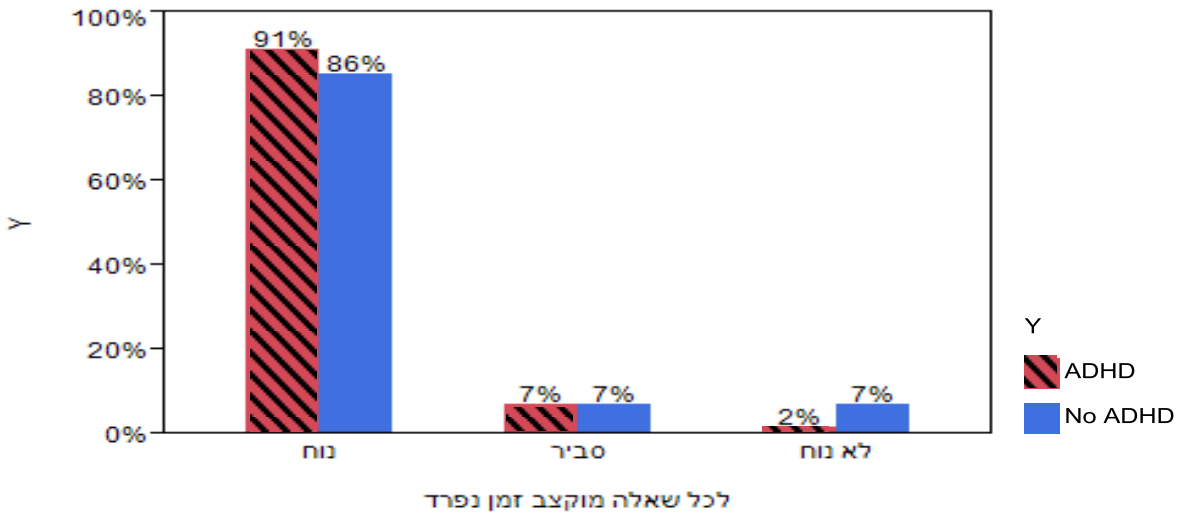
רוב גדול של הנבחנים (87%) חש שהצגת כל שאלה במסך נפרד היא נוחה, ומיעוט קטן בלבד מהם חש שהיא לא נוחה (2%). בניגוד למה שהיה אפשר לשער, נבחנים בעלי ADHD לא נבדלו מנבחנים רגילים במידת הנוחות שייחסו למאפיין זה (χ^2=0.88, p>0.05). ממצאים אלו מוצגים באיור 6.

איור 6 – התפלגות שביעות הרצון מהצגת כל שאלה במסך נפרד בבחינת מפע"ם



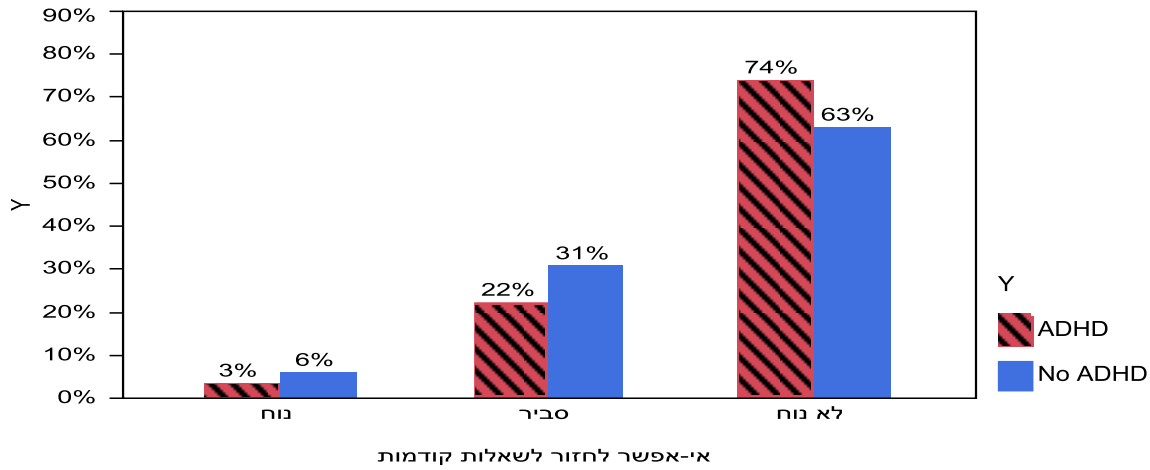
רוב גדול של הנבחנים (88%) חש גם שהקצבת הזמן בנפרד לכל שאלה היא מאפיין חיובי ונוח של השאלה. גם כאן לא נמצאו הבדלים מובהקים סטטיסטית בין הקבוצות (χ^2=2.53, p>0.05). ממצאים אלו מוצגים באיור 7.

איור 7 – התפלגות שביעות הרצון מהקצבת הזמן בנפרד לכל שאלה בבחינת מפע"ם



לבסוף, כצפוי, מרבית המשתתפים (67%) חשו שהעובדה שאי אפשר לחזור לשאלות קודמות היא מאפיין לא נוח של הבחינה, ורק מיעוטם חש שזהו מאפיין נוח שלה (5%). 28% מהנבחנים ציינו כי מאפיין זה סביר בעיניהם. שיעור הנבחנים בעלי ADHD שחשים שמאפיין זה אינו נוח עולה על שיעורם בקרב נבחנים ללא ADHD, אולם גם כאן ההבדל אינו מובהק סטטיסטית (2.04, χ2=p>0.05). ממצאים אלו מוצגים באיור 8.

איור 8 – התפלגות שביעות הרצון מהעובדה שאי אפשר לחזור לשאלות קודמות בבחינת מפע"ם

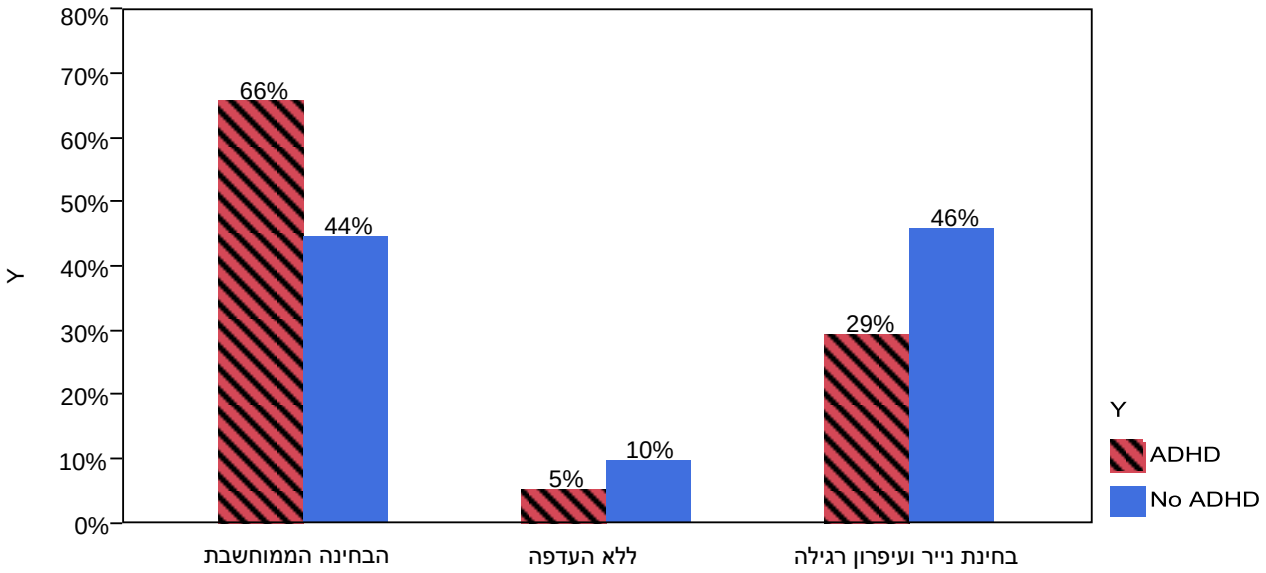


השוואה בין הבחינה הממוחשבת לבחינה הרגילה

הנבחנים התבקשו להשוות בין בחינת המפע"ם שבה נבחנו בניסוי לבין בחינת הנייר והעיפרון שהם מכירים ולציין איזו מהשתיים עדיפה בעיניהם. בפרט התבקשו הנבחנים לציין באיזו בחינה היו מעדיפים להיבחן ואם בבחינה האמתית שהם עתידים להיבחן בה הם חושבים שיצליחו יותר או פחות משהצליחו בבחינת הניסוי. כמו כן התבקשו הנבחנים להעריך מה תהיה מידת הצלחתם בכל אחד מהתחומים בכל אחת משתי הבחינות ולפרט עוד תחושות שלהם הקשורות לחוויית הבחינה (מידת החרדה, מידת הריכוז ועוד). תוצאות ניתוחים אלו מוצגות באיורים 9–10.

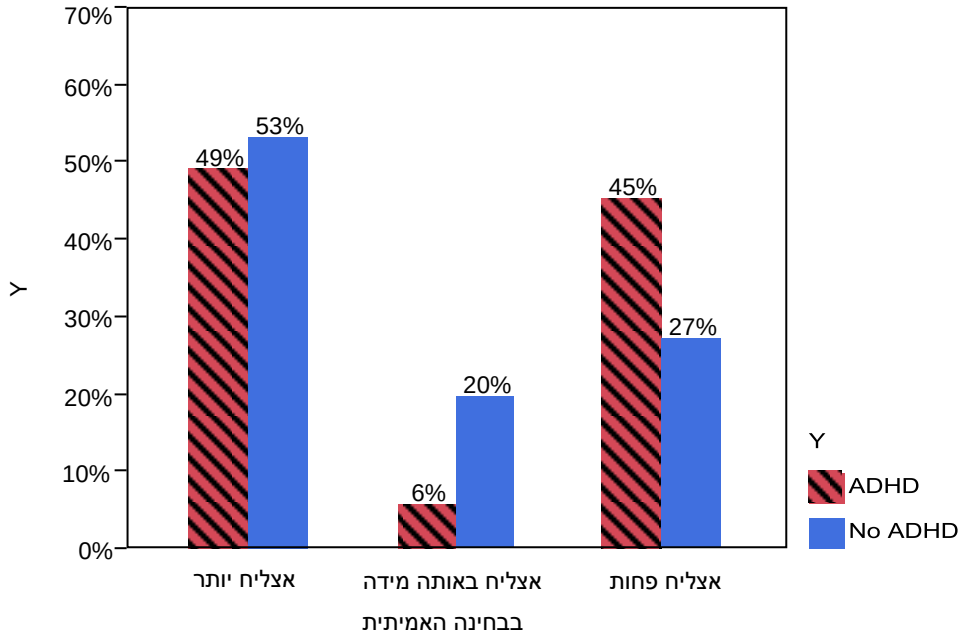
באופן כללי נמצא כי נבחנים בעלי ADHD נוטים להעדיף את הבחינה הממוחשבת על פני בחינת הנ"ע (ראה איור 9). נבחנים ללא ADHD אינם נוטים להעדיף אחת מאופנויות הבחינה באופן מיוחד, אך נבחנים בעלי ADHD מעדיפים בבירור את הבחינה הממוחשבת ($\chi 2=8.01$, $p<0.02$).

איור 9 – התפלגות העדפת הבחינה הממוחשבת על פני בחינת הנייר והעיפרון



כמעט מחצית מהנבחנים בעלי ADHD העריכו כי בבחינה האמתית יצליחו פחות מאשר בבחינת מפע"ם הניסויית. כמו כן, שיעור הנבחנים בעלי ADHD שסברו כי בבחינה האמתית יצליחו פחות היה גבוה במידה ניכרת משיעור הנבחנים ללא ADHD שסברו כך (45% לעומת 27% בהתאמה). עם זאת, שיעור דומה של משיבים בשתי הקבוצות העריך כי בבחינה האמתית יצליח יותר מאשר בבחינת הניסוי. בקרב נבחנים בעלי ADHD לא נמצא הבדל בין שיעור מי שהעריכו שהצליחו יותר בבחינת הניסוי לשיעור מי שהעריכו שיצליחו יותר בבחינה האמתית (χ^2=0.127, p>0.05), לעומת זאת, בקבוצת הנבחנים ללא ADHD היה רוב ברור לאלו שהעריכו שיצליחו יותר בבחינה האמתית (χ^2=6.91, p=0.03).

איור 10– התפלגות הערכת ההצלחה בבחינת הניסוי לעומת הבחינה האמתית



כדי לבדוק את הקשר שבין תפיסת חוויית הבחינה הממוחשבת (המידה שבה הנבחן מעדיף את הבחינה הממוחשבת או חושב שציונו בה יהיה גבוה יותר) לבין מידת השיפור בפועל (כלומר, מידת השיפור, אם היה, בציון הבחינה הממוחשבת על פני הציון בבחינת נייר ועיפרון), נבדקה מידת השיפור בפועל כפונקציה של תשובת הנבחן לשאלה איזו בחינה הוא מעדיף. ראשית נערך ניתוח שונות דו־גורמי שסוג הבחינה המועדפת וסוג הקבוצה היו בו משתנים בלתי תלויים, וההבדל שנמצא בין ציוני בחינת הניסוי לציוני הבחינה האמתית היה בו משתנה תלוי. מהניתוח עלה כי אין אפקטים מובהקים סטטיסטית לאף אחד מהמשתנים ($F(2, 139) = 0.45, P>.05$ ו־ $F(1, 139) = 0.36, p>.05$, בהתאמה) ולא לאינטראקציה ביניהם ($F(2, 139) = 0.54, p>.05$). פירוש הדבר הוא שלא נמצא קשר בין העדפתו של הנבחן באשר לסוג הבחינה שבו ייבחן לבין הפער שנמצא בין ציוניו בבחינה ממוחשבת ובין ציוניו בבחינת נייר ועיפרון.

כדי לבדוק את הקשר בין הערכת הנבחן את המידה שבה תעלה הצלחתו בבחינה הממוחשבת על הצלחתו בבחינה האמתית לבין מידה זו בפועל נערך עוד ניתוח השונות עם הצלחה המוערכת וסוג הקבוצה המשתנים הבלתי תלויים, והפער בין ציוני שתי הבחינות – המשתנה התלוי.

גם כאן לא נמצאו אפקטים מובהקים לאף אחד מהמשתנים ($F(2, 139) = 1.11, P > .05$ ו־ $F(1, 139) = 0.00, p>.05$, בהתאמה) ולא לאינטראקציה ביניהם ($F(2, 139) = 0.05, p>.05$). פירוש הדבר הוא שלא נמצא קשר בין הערכת הנבחן את המידה שבה תעלה הצלחתו באחת הבחינות על הצלחתו בבחינה האחרת לבין המידה שבה עלתה הצלחתו בפועל. ממוצעי הניתוחים כפונקציה של תשובות הנבחן מוצגים בלוח 12.

לוח 12 – ממוצעים וסטיות תקן של ההפרש בין הציונים, לפי העדפת הבחינה

	ציון כללי בבחינת הניסוי		ציון כללי בבחינה האמתית		ציון כללי בבחינת הניסוי פחות ציון כללי בבחינה האמתית	
אני מעדיף להיבחן:	ממוצע	ס"ת	ממוצע	ס"ת	ממוצע	ס"ת
בבחינה הממוחשבת	570.50	102.92	567.61	103.71	2.89	34.48
אין לי העדפה	592.91	64.75	597.00	62.73	-4.09	44.76
בבחינה פסיכומטרית רגילה	596.31	76.98	592.78	84.72	3.54	32.08
בבחינה הפסיכומטרית האמתית אצליח יותר	575.78	99.02	578.23	103.71	-2.45	33.77
בבחינה הפסיכומטרית האמתית אצליח באותה מידה	595.95	93.67	589.63	95.98	6.32	24.97
בבחינה הפסיכומטרית האמתית אצליח פחות	589.78	77.38	583.3	76.11	6.48	37.36

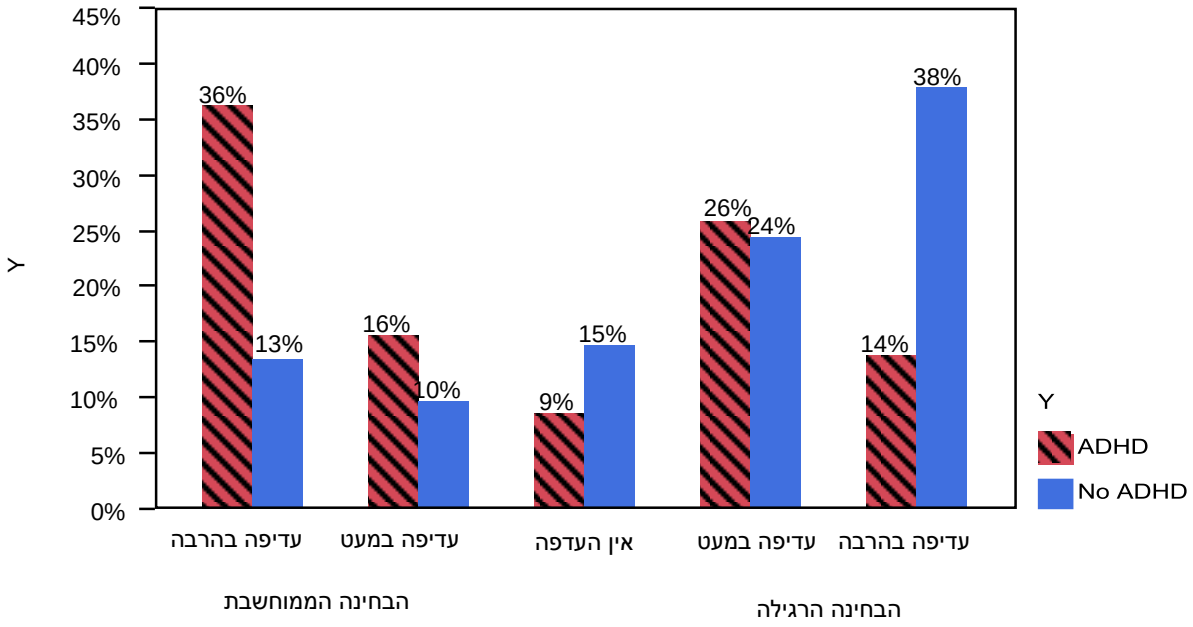
ממצאים אלו, המעידים על העדר קשר בין העדפת הנבחן ותפיסתו את הצלחתו להצלחתו בפועל, יידונו ביתר הרחבה בהמשך.

העדפת מאפיינים ספציפיים של הבחינה

הנבחנים התבקשו לציין איזו בחינה היו מעדיפים – ממוחשבת או נייר ועיפרון – בשמונה היבטים: פתרון השאלות בכל אחד משלושת התחומים, אופן סימון התשובות, מידת הריכוז בבחינה, מידת החרדה בבחינה, מספר השאלות בבחינה והקצאת הזמן בה. תוצאות הניתוחים מוצגות באיורים 11–18.

איור 11 מציג את התפלגות העדפת אחת הבחינות בתחום המילולי.

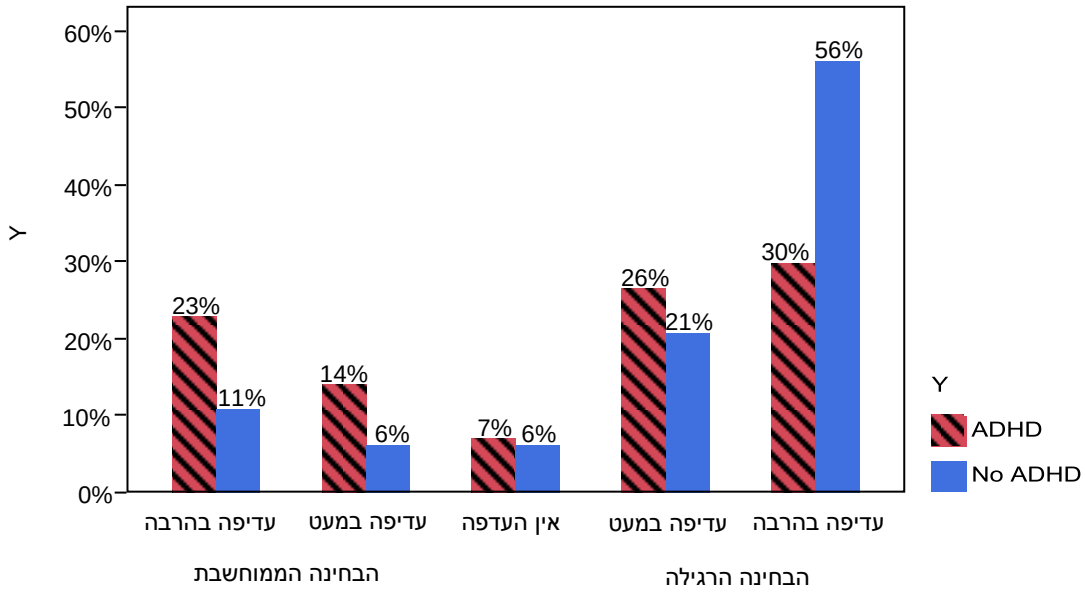
איור 11 – השוואה בין קבוצת הניסוי לקבוצת הביקורת בהעדפת אחת הבחינות (חשיבה מילולית)



מן האיור עולה כי מבחינת פתרון השאלות בתחום המילולי נבחנים בעלי ADHD מעדיפים באופן ברור את הבחינה הממוחשבת, ואילו נבחנים ללא ADHD מעדיפים בבירור את הבחינה הרגילה ($\chi^2=17.28$, $p<0.001$).

איור 12 מציג את התפלגות העדפת אחת הבחינות בתחום הכמותי.

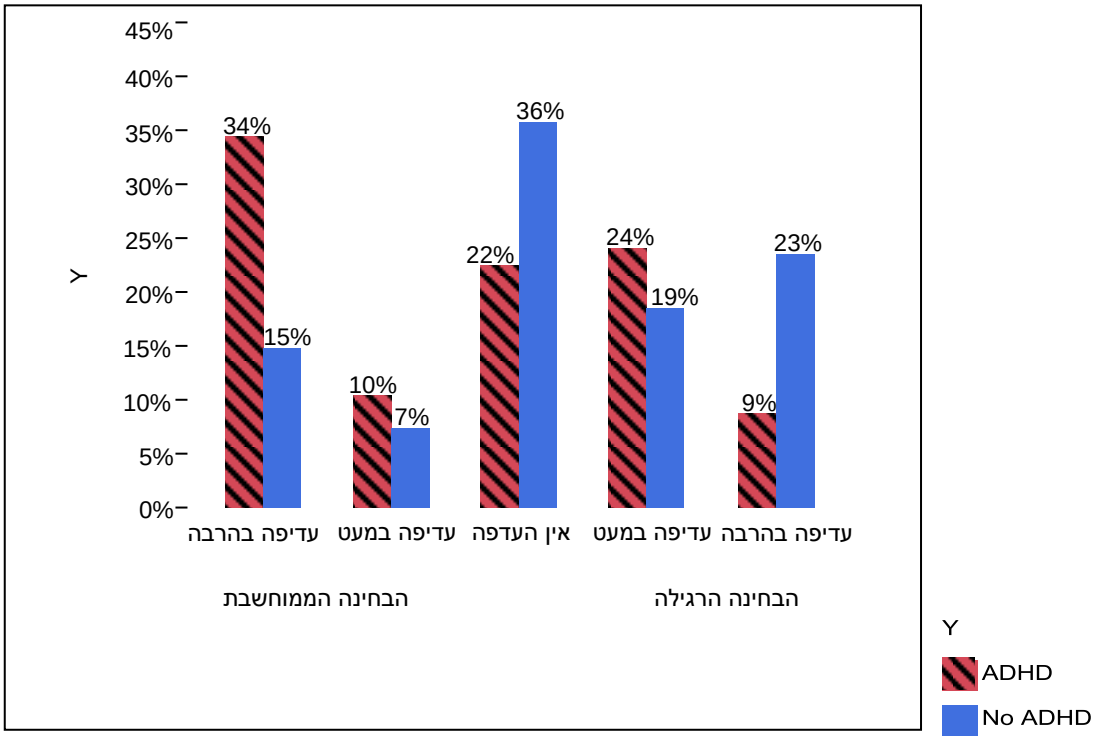
איור 12 – השוואת העדפת הבחינות אצל קבוצת הניסוי ואצל קבוצת הביקורת (חשיבה כמותית)



מן האיור עולה כי מבחינת פתרון השאלות בתחום הכמותי נבחנים ללא ADHD מעדיפים בבירור את הבחינה הרגילה, ואילו בקרב נבחנים בעלי ADHD אין העדפה ברורה לאחת הבחינות (χ^2=12.78, p<0.03).

איור 13 מציג את התפלגות העדפת אחת הבחינות בתחום האנגלית.

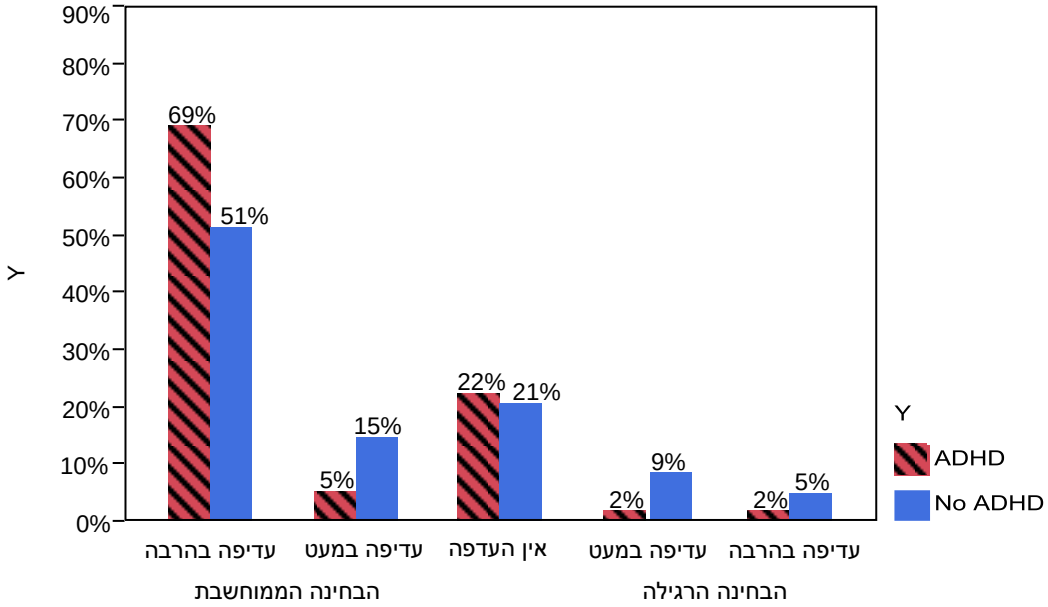
איור 13 – השוואת העדפת הבחינות אצל קבוצת הניסוי ואצל קבוצת הביקורת (אנגלית)



מן האיור עולה כי מבחינת פתרון השאלות בתחום האנגלית, בקרב נבחנים בעלי ADHD יש העדפה כלשהי לבחינה הממוחשבת, ואילו בקרב נבחנים ללא ADHD אין העדפה ברורה לאחת הבחינות (χ^2=13.2, p<0.02).

איור 14 מציג את התפלגות העדפת אחת הבחינות מבחינת אופן סימון התשובות.

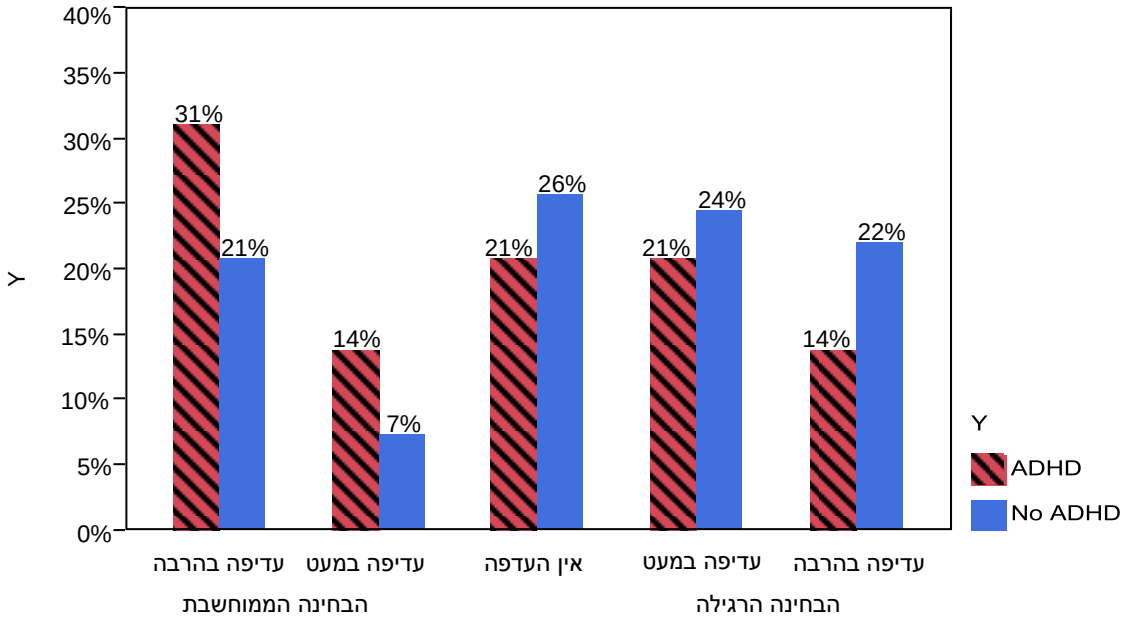
איור 14 – השוואת העדפת אופן סימון התשובות לשאלות אצל קבוצת הניסוי ואצל קבוצת הביקורת



מן האיור עולה כי שתי הקבוצות נוטות להעדיף את אופן סימון התשובות לשאלות בבחינה הממוחשבת, וההבדלים ביניהן קרובים למובהקות סטטיסטית (χ^2=9.22, p<0.06).

איור 15 מציג את התפלגות העדפת אחת הבחינות מבחינת רמת הריכוז הנדרשת מן הנבחן בבחינה.

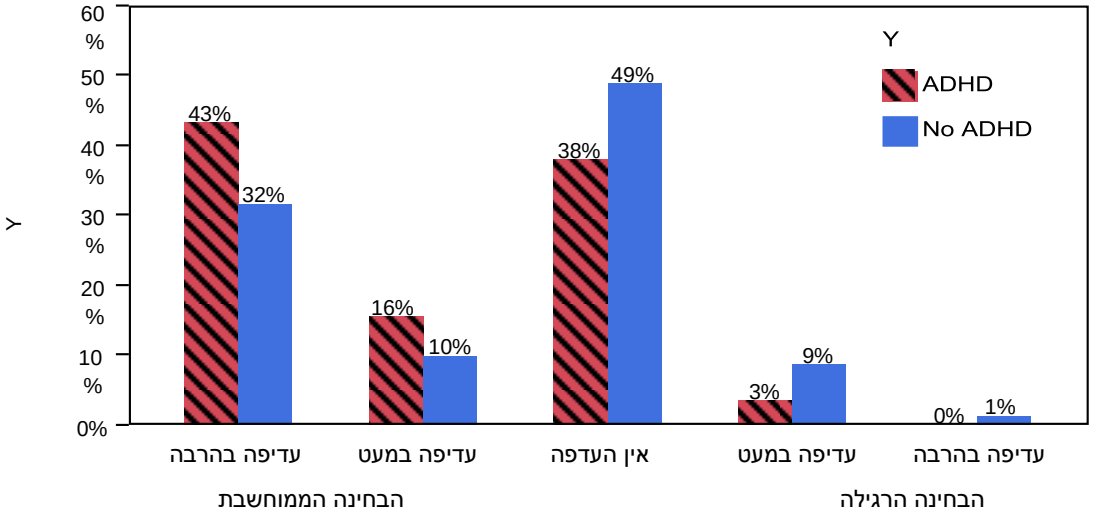
איור 15 – השוואת העדפת רמת הריכוז הנדרשת בבחינה אצל קבוצת הניסוי ואצל קבוצת הביקורת



מן האיור עולה כי מבחינת רמת הריכוז הנדרשת בבחינה, בקרב נבחנים בעלי ADHD יש העדפה קלה לבחינה הממוחשבת, ואילו בקרב נבחנים ללא ADHD אין העדפה ברורה לאחת הבחינות.

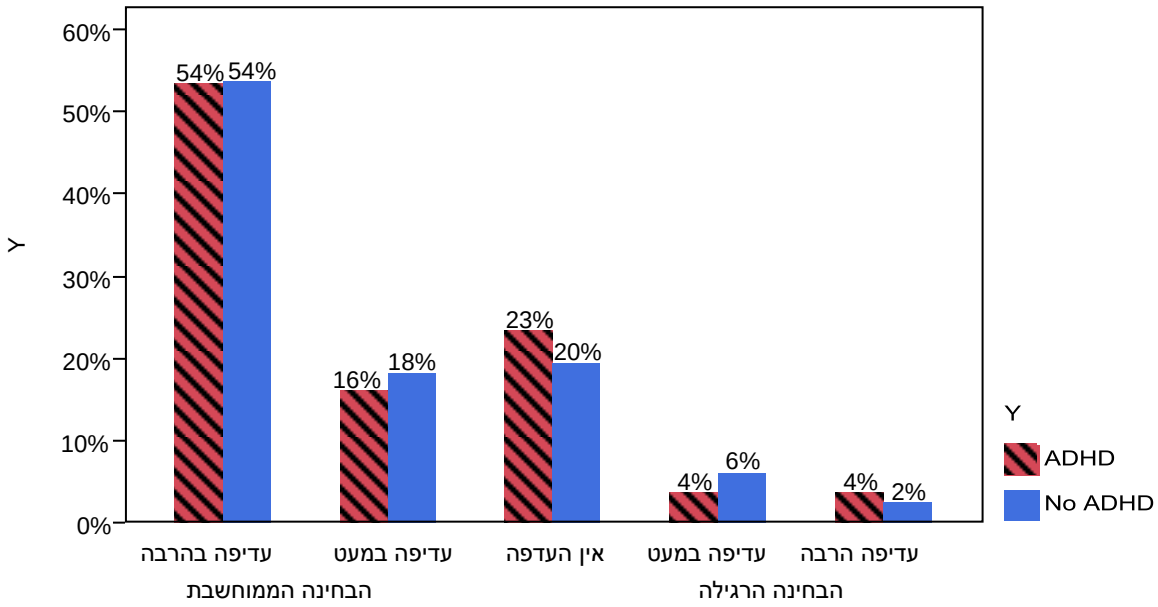
הבדלים אלו אינם מובהקים סטטיסטית. איור 16 מציג את התפלגות העדפת אחת הבחינות מבחינת רמת החרדה שהנבחן מעריך שיחוש במהלכה.

איור 16 - השוואת העדפת רמת החרדה המוערכת בבחינה בין קבוצת הניסוי לקבוצת הביקורת



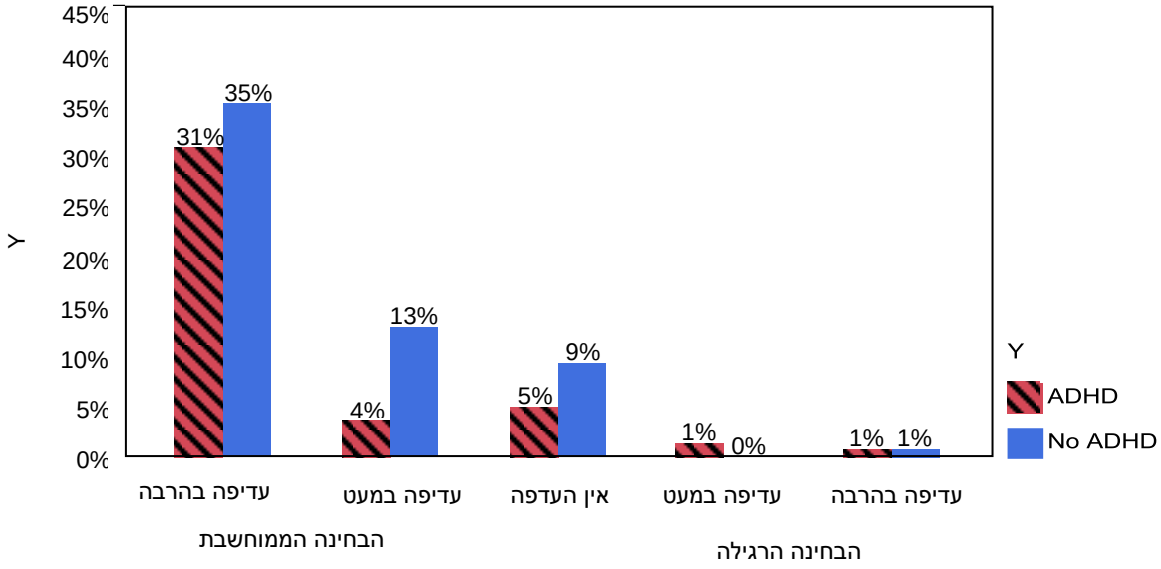
מן האיור עולה כי מבחינת החרדה בבחינה קבוצת הנבחנים בעלי ADHD נוטה להעדיף את הבחינה הממוחשבת, ואילו בקרב הנבחנים ללא ADHD אין העדפה ברורה לאחת הבחינות. ההבדלים בין הקבוצות אינם מובהקים סטטיסטית. איור 17 מציג את התפלגות העדפת אחת הבחינות מבחינת כמות הפריטים בבחינה.

איור 17 – השוואת העדפת מספר השאלות בבחינה אצל קבוצת הניסוי ואצל קבוצת הביקורת



גם כאן עולה העדפה ברורה של שתי הקבוצות כלפי הבחינה הממוחשבת. ההבדלים בין הקבוצות אינם מובהקים סטטיסטית. באיור 18 מוצגת התפלגות העדפת אחת הבחינות מבחינת זמן הבחינה הכולל.

איור 18 – השוואת העדפת זמן הבחינה הכולל אצל קבוצת הניסוי ואצל קבוצת הביקורת



מהאיור עולה כי שתי הקבוצות מעדיפות את זמן הבחינה הכולל בבחינה הממוחשבת, וההבדלים ביניהן אינם מובהקים סטטיסטית. לסיכום, מהממצאים עולה דפוס עקבי של הבדלים בין הקבוצות בכל הנוגע להעדפת אופנות הבחינה: נבחנים בעלי ADHD נוטים להעדיף את הבחינה הממוחשבת, וברוב הפרמטרים שנבדקו נבחנים ללא ADHD נוטים להעדיף את בחינת הנייר והעיפרון או שאינם מעדיפים אף אחת מהבחינות על פני האחרת.

דיון

מחקר זה נועד לבחון את יעילותה ואת תקפותה של הבחינה האדפטיבית הממוחשבת (מפע"ם) כהתאמה של הבחינה הפסיכומטרית לנבחנים בעלי ADHD. לבחינת מפע"ם כמה מאפיינים שבגללם היא עשויה להתאים לנבחנים בעלי מוגבלויות מגוונות: הקצאת זמן נדיבה, ריבוי הפסקות, הצגת שאלות באופן מאורגן ומובנה ומספר קטן של שאלות. במחקר זה השתמשנו בפרדיגמת "השיפור הדיפרנציאלי" (Sireci et al., 2005) כדי לבחון את תוקף השימוש בבחינה זו כהתאמה. על פי פרדיגמה זו, התאמה נחשבת תקפה אם השיפור שהיא מייצרת אצל נבחנים בעלי לקויות עולה על השיפור שהיא מייצרת אצל נבחנים ללא לקויות (Fuchs, Fuchs, Eaton, Hamlett, Binkley & Crouch, 2000). פרדיגמה זו מבוססת על ההנחה כי שיפור דיפרנציאלי מעיד על ספציפיות של הפיצוי: ההתאמה עוקפת קושי שהוא ספציפי ללקות, ולכן היא מביאה לידי שיפור קטן יותר בקרב כלל הנבחנים ללא המגבלה. ההשערה במחקר זה הייתה שבחינת המפע"ם תייצר שיפור דיפרנציאלי אצל נבחנים בעלי ADHD. לצורך בחינת ההשערה הושווה תפקודם של נבחנים בעלי ADHD בבחינת מפע"ם ובבחינת נייר ועיפרון לתפקודם של נבחנים ללא ADHD בבחינות אלו.

ממצאי המחקר לא תמכו בהשערת המחקר. אמנם נמצא כי בחינת המפע"ם שקולה לבחינת נ"ע לגבי נבחנים ללא ADHD, ואילו לגבי נבחנים עם ADHD היא מייצרת שיפור דיפרנציאלי קל, אך האינטראקציה בין סוג הקבוצה לרמת השיפור לא הייתה מובהקת סטטיסטית. ייתכן שממצאים אלו נובעים מבעיית עוצמה של מערך המחקר, שכלל קבוצות קטנות יחסית. כמו כן נמצאה שונות רבה בין הנבחנים בשיעור השיפור: תפקודם של כמה מהנבחנים השתפר מאוד, תפקודם של כמה מהם הורע מאוד, ותפקודם של כמה מהם לא השתנה. יתכן כי שונות רבה זו תרמה להנמכת העוצמה הסטטיסטית של המחקר ובכך תרמה לחוסר המובהקות הסטטיסטית של ממצאיו.

נוסף על כך, במחקר זה נמצא כי משך הזמן בו השתמשו נבחנים בעלי ADHD לפתרון כל סוגי השאלות בבחינה היה ארוך במובהק מן הזמן שנדרש לנבחנים ללא ADHD לשם כך. משמעותו של ממצא זה אינה ברורה לחלוטין. מחד, יתכן שנבחנים בעלי ADHD אמנם זקוקים לזמן רב יותר לעיבוד מידע כתוצאה מאיטיות הנגרמת בשל קשיי הקשב שלהם, והדבר מתבטא בקריאות חוזרות או בעיבוד כללי אטי יותר המושפע מליקויים בתהליכי קשב ובתהליכי זיכרון עבודה (Jacobson, Ryan, Martin, Ewen, Mostofsky, Denckla, & Mahone, 2011; Mayes & Calhoun, 2007). מאידך, יתכן שהזמן הנדרש להם לפתרון השאלות אינו שונה מהזמן הנדרש לנבחנים ללא ADHD, והם משתמשים בזמן הנותר למנוחה ולהפסקה, או שדעתם מוסחת מן הבחינה שלא בכוונה. אפשרות אחרת היא שכמה מהנבחנים ניחשו מהי מטרת הניסוי ופתרו את השאלות במשך זמן רב כדי להוכיח שהם אמנם זקוקים לזמן זה. מחקר זה אינו מאפשר להכריע בין שלושת ההסברים, וסביר כי כל אחד מהם עשוי להיות תקף לכל אחד מהנבחנים. מכל מקום, מממצאי המחקר לא עלה כי השימוש בזמן רב משפר את ההישגים, מה שמרמז שזמן זה אינו משמש לעיבוד נוסף, ואם השימוש בו נעשה לצורך הפסקות, הדבר אינו מייעל את תהליך העבודה.

ההשערה כי יימצאו הבדלים בין כמה מאפיינים של חישוב ציוני בחינת המפע״ם בקרב נבחנים בעלי ADHD למאפיינים אלו בקרב נבחנים ללא ADHD לא אוששה. אף על פי שנמצאו הבדלים באחוז ההתכנסות, במספר המתכנסים ובטעות התקן של המדידה בכיוון המשוער, היו הבדלים אלו קטנים יחסית ולא מובהקים סטטיסטית. גם לכך ייתכן שתרמה עוצמה סטטיסטית נמוכה.

אחד הממצאים שעלו מן הנתונים היה שונות ניכרת בתועלת שנבחנים שונים בעלי ADHD מפיקים מבחינת המפע״ם. אצל נבחנים מסוימים בעלי ADHD הניבה הבחינה שיפור גדול (כ־100 נקודות, שהן כסטיית תקן אחת), ואילו אצל אחרים לא הניבה שיפור כלל ולעתים אף פגעה בתפקודם (הישגיהם של כ־53% מהנבחנים בבחינת מפע״ם לא עלו על הישגיהם בבחינת ני״ע).

עם זאת, בקרב נבחנים ללא ADHD נמצא שיעור דומה של שיפור (54%). ייתכן שממצא זה מצביע על כך שדפוס השיפור אצל נבחנים בעלי ADHD דומה לדפוס השיפור אצל נבחנים ללא ADHD, וזאת שלא כמו אצל לקויי למידה מסוגים אחרים (למשל, לקויי קריאה), שלגביהם הראו מחקרים קודמים שיפור דיפרנציאלי גדול יותר ובשיעור גבוה יותר (Moshinskey & Kazin, 2004; לדני, 2015). כמו כן, השוואת התפלגות הפערים בקרב הנבחנים בעלי ADHD שנבחנו בניסוי בבחינת מפע״ם להתפלגות הפערים בקרב הנבחנים ללא ADHD שנבחנו בניסוי בבחינת מפע״ם העלתה כי אין הבדל בין שונויות ההתפלגות בשתי הקבוצות.

בתפקוד הכללי, הן בבחינת הניסוי והן בבחינה התפעולית, עלו השונויות בקרב נבחנים בעלי ADHD על השונויות בקרב נבחנים ללא ADHD. ממצא זה מתיישב עם הטענה, הנתמכת במחקרים, שלפיה בעלי ADHD מתאפיינים בהטרוגניות רבה (Newcorn, Halperin, Jensen, Abikoff, Arnold, Cantwell, & Vitiello, 2001) המתבטאת בתחומי תפקוד מגוונים. כמה מחקרים הבחינו בין דפוסים של תפקוד קוגניטיבי והתנהגותי (Fair, Bathula, Nikolas, & Nigg, 2012) בקרב בעלי ADHD, ובפרט בין דפוסים הנבדלים זה מזה במהירות עיבוד המידע: עיבוד אטי ומדויק לעומת עיבוד מהיר, אימפולסיבי ושגוי (Barkley, 2013; Becker, & Langberg, 2013; Hartman, Willcutt, Rhee, & Pennington, 2004; Lee, Burns, Snell, & McBurnett, 2014). הבחנה נוספת היא בין בעלי תסמינים של הפרעת קשב מהסוג הלא־קשוב לבעלי תסמינים של פעילות יתר (Fair, Bathula, Nikolas, & Nigg, 2012). חשוב לציין כי לעיתים קרובות דוחות האבחון המתקבלים ביחידת במו״ת אינם מאפשרים להבחין בין הסוגים.

השונות הרבה בין אנשים המאובחנים כבעלי ADHD נובעת גם מאבחנת יתר של ההפרעה, מהשונות בין שיטות האבחון ובין המאבחנים ומכך שהאבחנה ניתנת על סמך הערכה סימפטומטית ולא אטיולוגית. יתכן כי גם ההבדלים בין דפוסי תפקוד תורמים להבדלים התפקודיים הנצפים אצל המאובחנים ולהבדלים בשיפור אצלם בבחינת מפע״ם.

לבסוף, פעמים רבות נטען כי מצבים מאתגרים ומהירים, שהגירויים מתחלפים בהם בקצב מהיר ואינם משרים שעמום, משפרים את תפקודם של בעלי ADHD. המצדדים בטענה זו מאמינים שמצבים אלו משפרים את גיוס הקשב בקרב הסובלים מההפרעה (Barkley, 2007). אם טענה זו

נכונה, ביצועיהם של כמה מהנבחנים בבחינת נייר ועיפרון אמורים לעלות על ביצועיהם בבחינת מפע״ם, משום שבחינה זו מתנהלת בקצב מהיר יותר מבחינת מפע״ם. עניין זה יידון לעומק להלן.

מאפייני הבחינות והשפעתם האפשרית על התפקוד

קיימים הבדלים רבים בין בחינת מפע״ם לבחינת נייר ועיפרון, החל מפורמט הצגת השאלות, המשך בהקצאת הזמן וכלה במאפייני חדר הבחינה ובאופן העברת הבחינה. כמו כן, בכל בחינה יש כמה תחומים (חשיבה מילולית, כמותית ואנגלית) הדורשים כישורים שונים זה מזה, והדבר עשוי להביא לידי תפקוד לא אחיד של הנבחן בתוך הבחינה.

אחד ההיבטים שבהם ההבדל בין הבחינות בולט במיוחד הוא קצב הבחינה. בחינת מפע״ם מתנהלת בקצב איטי הרבה יותר, וזאת בשל הקצבת הזמן הנדיבה לנבחנים בה (כל שאלה בבחינה זו מוצגת בנפרד, ומוקצב לה זמן הארוך עד פי ארבעה מהזמן הממוצע המוקצה לשאלה בבחינת נייר ועיפרון). סביר שתפקודו של כל נבחן הושפע מכך באופן אחר ושיש נבחנים שהקצב האטי של הבחינה הוא בעוכריהם. בפרט, יש נבחנים בעלי ADHD שתנאים מסוג זה גורמים להם דווקא לאבדן ריכוז, לחלומות בהקיץ ולעייפות, שלא כתנאיה של בחינה מאתגרת, משתנה במהירות ומוגבלת בזמן כמו בחינת נייר ועיפרון. כמו כן, בחינת מפע״ם יכולה להימשך עד 5.5 שעות וכך לגרום לעייפות ולקושי בשמירה על קשב מתמשך – שתיים מהבעיות האופייניות ביותר לבעלי ADHD. יש לציין שאת הקצב בבחינת מפע״ם הנבחן עצמו קובע – הוא אינו חייב להתנהל באטיות, אך יכול לבחור לעשות כך. כלומר, ייתכן שבחינה זו מאפשרת לנבחן (שלעיתים קרובות מתקשה לנהל זמן כראוי בגלל ה־ADHD) להתנהל בדרך שאינה יעילה מבחינתו.

מאידך, בחינת מפע״ם היא בחינה מובנית יותר מבחינת נייר ועיפרון. כאמור, בבחינה זו המחשב בוחר את השאלות ומציג אותן אחת אחת, וכך דעתו של הנבחן אינה מוסחת בגלל שאלות אחרות או בגלל טקסט שאינו רלוונטי לשאלה שהוא מנסה לפתור באותו רגע. כמו כן, כל שאלה מוצגת בכתב גדול ומאחר שהנבחן מזין את התשובות ישירות למחשב, הוא אינו נדרש לעבור בין חוברת כבחינה לבין גיליון כתשובות, ולכן אפשר להניח שהסיכוי לתופעות כגון הסטה בסימון התשובות, בלבול בין שאלות וכדומה – תופעות העלולות להחריף בשל קשיי קשב – קטֵן.

ייתכן שממצאי המחקר משקפים קיזוז של יתרונות בחינת המפע״ם ביחס לנבחנים בעלי ADHD עם חסרונותיה ביחס אליהם, ואולי בשל כך הממצאים אינם מצביעים על שיפור גורף בציוניהם של הנבחנים בעלי ADHD. נראה כי יש לנסות לבדוק כל נבחן בנפרד על פי מאפייניו ותפקודו, וכך להחליט אם הבחינה הממוחשבת צפויה לשפר את הישגיו. סוגיה זו תידון ביתר הרחבה להלן, תוך התייחסות לשיפוטי המומחים.

שיפוטי מומחים

בשל השונות הגבוהה בשיפור שהושג בקרב נבחנים בעלי ADHD בבחינת מפע״ם, והיות שכמה מהם קיבלו בה ציון נמוך מזה שקיבלו בבחינת נייר ועיפרון, היה צורך לערוך ניתוח המשך כדי לנסות לקבוע אם אפשר לאתר מראש את הנבחנים שציונם בבחינה ממוחשבת עתיד לעלות על ציונם בבחינת ני״ע על סמך נתוני האבחון בתיקם. לשם כך ערכו שלושה מומחים בלתי תלויים שיפוטים לכל נבחן – שיפוטים שנועדו לחזות אם ציונו של נבחן זה בבחינה ממוחשבת יעלה על

ציונו בבחינת נ״ע, וזאת על פי ממדים שפורטו בפרק ״שיטת המחקר״. נמצא כי שיפוטי המומחים, שנערכו על סמך נתונים תפקודיים שמצאו בתיקי בקשת ההתאמות של הפונים (אבחונים דידקטיים ופסיכולוגיים, אבחוני קשב, חו״ד מורים וחומרי רקע שונים), הצליחו להבחין בין הקבוצות. כמו כן נמצא כי ממוצע השיפור בקרב הקבוצה שלגביה קבעו המומחים כי ציונה ישתפר בבחינת מפע״ם היה גבוה מזה של הקבוצה שלגביה קבעו המומחים כי ציונה לא ישתפר בבחינת מפע״ם.

כפי שעלה מהתפלגות הציונים, הקבוצה שלגביה קבעו השופטים שציוניה צפויים להשתפר הייתה מורכבת משתי תת-קבוצות מובחנות – קבוצה אחת שציונה הממוצע השתפר מאוד וקבוצה אחרת שציונה הממוצע הורע. בקבוצה שלגביה קבעו השופטים שציוניה אינם צפויים להשתפר הייתה התפלגות הציונים רציפה יותר, אם כי גם שם נמצאו שינויים משני הסוגים. נראה שתופעה זו ראויה לחקירה מעמיקה יותר. בפרט יש לבחון לעומק אם יש מאפיינים המשותפים לחברי הקבוצה שחזו השופטים כי ציוניה ישתפרו אך בפועל הורעו, ולבחון את השפעת מאפיינים אלו – הן תאורטית והן אמפירית – על האופן שבו הם משפיעים על עמדתו של שופט קליני מנוסה לעומת האופן שבו הם משפיעים על תפקוד הנבחנים בבחינה. למשל, יש לבדוק אם קיימים משתנים המשפיעים על שיפוט המומחים לחיוב (המומחים מאמינים שציון הנבחן ישתפר בבחינת מפע״ם) אך לעומת זאת משפיעים על מגמת השיפור לרעה (בפועל חלה ירידה בציון).

לבסוף, חשוב לציין כי בדיקה זו הייתה ראשונית ומצומצמת, ומתוצאותיה אי אפשר להסיק מסקנות גורפות לגבי המדיניות הראויה לקביעת התאמות. כדי לפתח אמצעי אבחון שיאפשר לקורא מומחה לקבוע אם פונה כלשהו צפוי להצליח בבחינת מפע״ם יותר מבבחינת נייר ועיפרון יש צורך לחזור על הבדיקה בהקפדה על קריטריוני החלטה ברורים ומוגדרים היטב, לחקור את השופטים לגבי תהליך קביעת השיפוט, לבדוק את מהימנות השיפוטים, להגדיל את המדגם במידה ניכרת ולקבוע את תוקף הניבוי של תוצאות ההערכה. בדיקה כזו חורגת מהיקף מחקר זה.

ממצאי שאלון המשוב

שאלון המשוב הועבר לכל הנבחנים בבחינת מפע״ם, בעלי ADHD ושאינם בעלי ADHD. מטרתו הייתה להשיג משוב לגבי חוויית הבחינה, נגישותה לנבחן ונוחות פתירתה, ולהשוות בעניין זה בין נבחנים בעלי ADHD לנבחנים ללא ADHD. באופן כללי הראו הממצאים כי הבחינה הממוחשבת נתפסת כבחינה נוחה, ידידותית ונגישה בקרב שתי הקבוצות.

בתחום הכמותי נמצא כי מרבית הנבחנים מעדיפים את בחינת הנייר והעיפרון. הסבר אפשרי לממצא זה הוא שבתחום הכמותי שאלות רבות מבוססות על איורים ותרשימים, ופתרון שאלות כאלו במחשב, מצב שבו נבחנים רבים מבקשים להעתיק את התרשים לדפי הטיוטה, מורכב יותר. בתחום החשיבה המילולית עולה אינטראקציה ברורה בין סוג הקבוצה לבין סוג הבחינה – נבחנים בעלי ADHD מעדיפים בבירור את הבחינה הממוחשבת, ואילו נבחנים ללא ADHD מעדיפים בבירור את בחינת הנייר והעיפרון. בתחום האנגלית ניכרת מגמה דומה, אם כי חלשה יותר. כאמור, לא נמצאו הבדלים מובהקים בין תפקודי הנבחנים בשתי אופנויות הבחינה.

מהממצאים עולה דפוס עקבי יחסית של הבדלים בין העדפותיהן של שתי הקבוצות לגבי שתי הבחינות. נבחנים בעלי ADHD נוטים להעדיף את הבחינה הממוחשבת ברוב הפרמטרים שנבדקו.

לעומתם, נבחנים ללא ADHD מגלים העדפה לבחינת הנייר והעיפרון או שאינם מגלים העדפה כלל. יתכן כי ממצא זה קשור לכך שנבחנים בעלי ADHD, אשר הגישו בקשה למדור תנאים מותאמים במרכז הארצי לבחינות ולהערכה, מודעים למאפייניה של בחינת מפע"ם, ובפרט להקצאת הזמן הנדיבה יותר בבחינה זו. ייתכן שהערכתם לגבי נוחות הבחינה אינה נובעת מהתנסותם בה בפועל, אלא מתפיסתם האפריורית שזוהי אופנות הבחינה העדיפה להם.

כאמור, נמצא כי נבחנים בעלי ADHD העדיפו את הבחינה הממוחשבת יותר מנבחנים ללא ADHD. עם זאת, ממצא מעניין העולה מן הנתונים הוא העדר קשר בין הערכת הנבחנים (בעלי הפרעת ADHD או בלעדיה) לגבי השיפור בביצועיהם בבחינת מפע"ם ("אני מעדיף את הבחינה הממוחשבת" ו"בבחינה הממוחשבת אצליח יותר") לבין מידת השיפור שהראו בפועל. את השאלון מילאו הנבחנים **לאחר** שנבחנו בבחינה הממוחשבת ולפני שקיבלו את הציון, ולכן התבססה הערכתם לא רק על ציפיותיהם מהבחינה הממוחשבת, אלא גם ובעיקר על התנסותם בה בפועל. נתון זה מחזק את המסקנה כי הקשר בין חווייתו של הנבחן ומידת הערכתו באשר להצלחתו לבין המציאות הוא קלוש. ממצא זה מתיישב עם ממצאי מחקרים קודמים, אשר הראו כי תפיסתו הסובייקטיבית של הנבחן אינה מנבאת את הצלחתו בבחינה עם התאמות לעומת הצלחתו בבחינה ללא התאמות, מעבר לרמת הניחוש (Elbaum, Arguelles, Campbell, & Saleh-Bardawil, 2004). הסתייגות אחת ממסקנה זו קשורה לעובדה כי קיים confounding בין היותה של הבחינה בחינת נייר ועיפרון לבין היותה בחינה תפעולית אמתית, שהציון בה משפיע על עתיד הנבחנים, ולא בחינה לצורכי ניסוי בלבד – עובדה שלא הובהרה בשאלון כראוי.

מגבלות המחקר

לפני הדיון בהשלכות המחקר יש להכיר בכמה ממגבלותיו. ראשית, מחקר זה סובל ממגבלת עוצמה. בשל הקושי לגייס נבחנים לבחינות מקדימות, ובפרט נבחנים בעלי ADHD (קבוצה קטנה הנוטה להתאפיין במוטיבציה נמוכה, בקשיי ארגון ובקשיי עמידה במחויבות), היו הקבוצות במחקר זה קטנות יחסית. אין ספק שגודל המדגם הקטן פגע בעוצמה הסטטיסטית, והדבר התבטא בכך שפערים רבים נמצאו בכיוון המשוער אך לא הגיעו למובהקות סטטיסטית. בשל שינוי פורמט הבחינה הפסיכומטרית ביולי 2012 לא התאפשר לגייס נבחנים נוספים למחקר ולהגדיל את הקבוצות.

נוסף על כך, במחקר זה ניתנה בחינת מפע"ם כ"חבילה" אחת הכוללת תוספת זמן, הפסקות והבנייה מאורגנת ומסודרת של הבחינה. מבנה בחינת מפע"ם אינו מאפשר חקירה של כל אחד ממרכיביה בנפרד ובדיקה של תרומתו להישגי הנבחן.

השלכות המחקר

השלכות על מחקרי המשך

ממצאי המחקר מעלים לדיון מחודש שאלות הנוגעות למנגנונים הקוגניטיביים הפגועים ב-ADHD ולדרכי הפיצוי הראויות על פגיעות אלו. כאמור, מחקר זה לא בדק כל מרכיב של בחינת המפע"ם בנפרד, ולכן אי אפשר לקבוע אילו מהמרכיבים עשויים לתרום לתפקודם של נבחנים

בעלי ADHD. סוגיה זו קשורה לשאלה מהו המנגנון או המנגנונים הפגועים ב-ADHD: האם זוהי איטיות בתפקוד, המשפיעה על תוצאות הבחינה? האם מדובר בהתעייפות, הדורשת התרעננות והפסקות? האם אלו הם קשיי ארגון וצורך לעבור תדיר מחוברת הבחינה אל גיליון התשובות וחוזר חלילה? נוסף על כך, נשאלת השאלה אם כותבי המבחן והמוסדות האקדמיים מעוניינים לעקוף את השפעות ה־ADHD על הישגי הנבחנים, או שהן חלק מהפרמטרים שהם בודקים. הבדיקות שנערכו במחקר זה אינן יכולות לספק תשובות לשאלות אלו, והן נותרות פתוחות למחקר נוסף. בחינת מנגנוני הפיצוי הקיימים והבנת פעולתם יכולות לא רק לשפוך אור על סוגיות הקשורות לדרכי היבחנות ולמתן התאמות, אלא גם לתרום להבנה מעמיקה יותר של המנגנונים הקוגניטיביים הפגועים ב־ADHD. זאת באמצעות בדיקת השערות הנוגעות לעקיפת מנגנונים קוגניטיביים אצל בעלי ADHD ובחינת רמת השיפור בתפקודם כפונקציה של מנגנון הפיצוי.

השלכות יישומיות

למחקר זה יש כמה השלכות יישומיות בתחום קביעת ההתאמות לנבחנים בעלי ADHD. חשוב לזכור כי כל תהליך קביעת התאמות הוא פרטני בהכרח, וזאת משום שהוא מבוסס על מאפייני הבחינה, על מאפייני הנבחן ועל המידה שבה אפשר לפצותו בלי לפגוע במדידה. לעומת זאת, תוצאות מחקר אמפירי מעלות כמעט תמיד ממצאים ממוצעים ברמה הקבוצתית, ולכן לא נכון לגזור מהן החלטות מעשיות לגבי נבחן ספציפי.

ממצאי המחקר מראים שיש קבוצת נבחנים בעלי ADHD שציונם בבחינת מפע״ם עולה על ציונם בבחינת נייר ועיפרון ויש קבוצת נבחנים שציונם בבחינת מפע״ם זהה לציונם בבחינת נייר ועיפרון או נמוך ממנה. כמו כן, ממצאי המחקר מצביעים על כך שקוראים מומחים, בעלי הכשרה וניסיון בקביעת התאמות לבעלי לקויות ממינים שונים, יכולים להבחין בין הקבוצות במידה כלשהי.

כדי לשפר את רמת הדיוק של השיפוטים המקצועיים המנבאים את מידת השיפור בהישגים בבחינה ממוחשבת, על השופטים יהיה לבדוק את מהירות עיבוד המידע והתפקוד של הנבחן, את רמת הארגון שלו, את רמת המוסחות שלו ואת הצורך שלו בחזרות על הוראות ושאלות, ועל סמך תוצאות הבדיקה לקבוע, בחלק מן המקרים, אם בחינת מפע״ם היא בגדר התאמה שהוא זקוק לה. מובן שגם בהליך כזה, כמו בכל הליך שיפוט סובייקטיבי, עלולות להתגלע טעויות, וזאת גם אם הוא מבוסס על נתונים כמותיים ומושווים לנורמה. בטווח הארוך יש ללוות הליכים חדשים מסוג זה במחקר הוגנות, אשר יבדוק את תוקף הניבוי מול ציון הקריטריון.

במסגרת הדיון הנוגע לפיצוי נבחנים בעלי ADHD ועקיפת ההפרעה לתפקודם בבחינה יש לדון לעומק בשאלה עד כמה הפרעה זו רלוונטית למשתנים הנמדדים בבחינה, כלומר לתפקודם האקדמי. במילים אחרות, האם ציון בבחינה שהותאמה לנבחן כדי למקסם את תפקודו ינבא את תפקודו האקדמי בעתיד (ניבוי שהוא מטרתה של הבחינה הפסיכומטרית) טוב יותר מציון שהתקבל בבחינת נ״ע? סוגייה זו רלוונטית במיוחד לתפקודם של מועמדים בעלי ADHD, שכן תפקודם מתאפיין בתנודות ושינויים בתלות בהקשר התפקוד, והתאמה המעניקה להם תנאים מיטביים עלולה להסוות את הבעיה בתפקודם. על כן נשאלת השאלה אם נכון לאפשר לנבחנים אלו להיבחן בתנאים העשויים להסוות בעיה העתידה לבוא לידי ביטוי במהלך לימודיהם האקדמיים.

מקורות

אורן, כ., קנת-כהן, ת., ברונר, ש. (ספטמבר 2007). נתונים מקובצים על תוקף מערכת המיון לאוניברסיטאות בחיזוי הצלחה בלימודי שנה א' (מחזורים תשס"ג - תשס"ה). (דוח מחקר מס' 342). המרכז הארצי לבחינות: ירושלים.

אנגל, א. (יולי, 2008). השפעת תוספת זמן בבחינה הפסיכומטרית על נבחנים עם ADHD. (דוח מחקר מס' 351). המרכז הארצי לבחינות: ירושלים.

לדני, ר. (2015). יעילות ותקפות בחינת המפע"ם כהתאמה פוטנציאלית ללקויי קריאה. עבודת גמר לתואר מוסמך: האוניברסיטה העברית.

לוינסקי, י. (2014, 9 באפריל). מהי הפרעת קשב אמתית, ומדוע אינה מטופלת נכון. *הארץ (סופשבוע)*. נדלה מ-http://www.haaretz.co.il/magazine/ayelet-shani/.premium-1.2293570 ב-26.5, 2014.

קלפר, ד. טורוול, א., קנת-כהן, ת., ואורן, כ (2013, ינואר). הוגנות מערכת המיון להשכלה הגבוהה כלפי מבקשי תנאים מותאמים בבחינה הפסיכומטרית. (דוח מחקר מס' 386). המרכז הארצי לבחינות: ירושלים.

קנת-כהן, ת., ברונר, ש. ואיתן, מ. (2008, אפריל). השפעת תוספת זמן על הציונים ועל מדרג הנבחנים בשלושת תחומי מכפ"ל. (דוח מחקר מס' 344). המרכז הארצי לבחינות: ירושלים.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders. Arlington, VA: American Psychiatric Publishing.

Amrein, A. L., & Berliner, D. C. (2002). High-stakes testing & student learning. Education Policy Analysis Archives, 10, 18.

Baker, Frank (2001). The Basics of Item Response Theory. ERIC Clearinghouse on Assessment and Evaluation, University of Maryland.

Barkley, R. A. (1998). ADHD in children and adolescents: Theory, diagnosis, and management. Developing Child XIX Conference, Champaign, IL.

Barkley, R. A. (2007). Attention-Deficit/Hyperactivity Disorder. In: D. A. Wolfe, & E. J. Mash. (Eds). *Behavioral and Emotional Disorders of Adolescents: Nature, Assessment, and Treatment* (pp. 91-152). New York: The Guilford Press.

Barkley, R. A. (2013). Distinguishing sluggish cognitive tempo from ADHD in children and adolescents: executive functioning, impairment, and comorbidity. *Journal of Clinical Child & Adolescent Psychology*, *42*(2), 161-173.

Becker, S. P., & Langberg, J. M. (2013). Sluggish Cognitive Tempo Among Young Adolescents With ADHD Relations to Mental Health, Academic, and Social Functioning. *Journal of Attention Disorders*, *17*(8), 681-689.

Biederman, J., Faraone, S. V., Doyle, A., Lehman, B. K., Kraus, I., Perrin, J., & Tsuang, M. T. (1993). Convergence of the Child Behavior Checklist with Structured Interview-based Psychiatric Diagnoses of ADHD children with and without comorbidity. *Journal of Child Psychology and Psychiatry*, *34*(7), 1241-1251.

Brown, T. E., Reichel, P. C., & Quinlan, D. M. (2011). Extended time improves reading comprehension test scores for adolescents with ADHD. *Open Journal of Psychiatry*, *1*(3), 79.

Cahan, S., Alkoby, M. & Nirel, R. (2010). The Extra-Examination Time Granting Policy: A Reconceptualization. 2010 Annual Meeting of the American Educational Research Association (AREA), Denver.

Cohen, Y., Ben-Simon, A., Moshinsky, A, & Eitan, M. (July, 2008). Computer-Based Testing (CBT) in the service of test accommodations. (RR no. 348). National Institute for Testing and Evaluation: Jerusalem.

Chhabildas, N., Pennington, B. F., & Willcutt, E. G. (2001). A comparison of the neuropsychological profiles of the DSM-IV subtypes of ADHD. *Journal of Abnormal Child Psychology*, *29*(6), 529-540.

Elbaum, B., Arguelles, M. A., Campbell, Y., & Saleh-Bardawil, M. (2004). Effects of a student-reads-aloud accommodation on the performance of students with and without learning disabilities on a test of reading comprehension. *Exceptionality*, *12*, 71-87.

Fair, D. A., Bathula, D., Nikolas, M. A., & Nigg, J. T. (2012). Distinct neuropsychological subgroups in typically developing youth inform heterogeneity in children with ADHD. *Proceedings of the National Academy of Sciences*, *109*(17), 6769-6774.

Fuchs, L. S., Fuchs, D., Eaton, S. B., Hamlett, C. Binkley, E., & Crouch, R. (2000). Using objective data sources to enhance teacher judgments about test accommodations. *Exceptional Children*, *67*(1), 2000, 67-81.

Gregg, N., & Nelson, J. M. (2012). Meta-analysis on the effectiveness of extra time as a test accommodation for transitioning adolescents with learning disabilities more questions than answers. *Journal of learning disabilities*, *45*(2), 128-138.

Harrison, A. G., Edwards, M. J., & Parker, K. C. (2007). Identifying students faking ADHD: Preliminary findings and strategies for detection. *Archives of Clinical Neuropsychology*, *22*(5), 577-588.

Holdnack, J. A., Moberg, P. J., Arnold, S. E., Gur, R. C., & Gur, R. E. (1995). Speed of processing and verbal learning deficits in adults diagnosed with attention deficit disorder. *Cognitive and Behavioral Neurology*, *8*(4), 282-292.

Hartman, C. A., Willcutt, E. G., Rhee, S. H., & Pennington, B. F. (2004). The relation between sluggish cognitive tempo and DSM-IV ADHD. *Journal of abnormal child psychology*, *32*(5), 491-503.

Jacobson, L. A., Ryan, M., Martin, R. B., Ewen, J., Mostofsky, S. H., Denckla, M. B., & Mahone, E. M. (2011). Working memory influences processing speed and reading fluency in ADHD. *Child Neuropsychology*, *17*(3), 209-224.

Lee, S., Burns, G. L., Snell, J., & McBurnett, K. (2014). Validity of the sluggish cognitive tempo symptom dimension in children: Sluggish cognitive tempo and ADHD-inattention as distinct symptom dimensions. *Journal of abnormal child psychology*, *42*(1), 7-19.

Lewandowski, L. J., Lovett, B. J., Parolin, R., Gordon, M., & Codding, R. S. (2007). Extended time accommodations and the mathematics performance of students with and without ADHD. *Journal of Psychoeducational Assessment, 25(1), 17-28.*

Jasinski, L. J., & Ranseen, J. D. (2011). Malingered adHd evaluations: a further complication for accommodations reviews. *The Bar Examiner, 79,* 6-16.

Mayes, S. D., & Calhoun, S. L. (2007). Learning, attention, writing, and processing speed in typical children and children with ADHD, autism, anxiety, depression, and oppositional-defiant disorder. *Child Neuropsychology*, *13*(6), 469-493.

Moshinsky, A., & Kazin, C. (2004). Constructing a computerized psychometric adaptive test for university applicants with disabilities. National Institute for Testing & Evaluation.

Newcorn, J. H., Halperin, J. M., Jensen, P. S., Abikoff, H. B., Arnold, L. E., Cantwell, D. P., ... & Vitiello, B. (2001). Symptom profiles in children with ADHD: effects of comorbidity and gender. *Journal of the American Academy of Child & Adolescent Psychiatry*, *40*(2), 137-146.

Pariseau, M. E., Fabiano, G. A., Massetti, G. M., Hart, K. C., & Pelham Jr, W. E. (2010). Extended time on academic assignments: Does increased time lead to improved performance for children with attention-deficit/hyperactivity disorder? *School Psychology Quarterly*, *25*(4), 236.

Phillips, S. E. (1994). High-stakes testing accommodations: Validity versus disabled rights. *Applied Measurement in Education*, *7*, 93-120.

Rucklidge, J. J., & Tannock, R. (2002). Neuropsychological profiles of adolescents with ADHD: Effects of reading difficulties and gender. *Journal of child psychology and psychiatry*, *43*(8), 988-1003.

Schoechlin, C., & Engel, R. R. (2005). Neuropsychological performance in adult attention-deficit hyperactivity disorder: meta-analysis of empirical data. *Archives of Clinical Neuropsychology*, *20*(6), 727-744.

Sireci, S. G., Scarpati, S. E., & Li, S. (2005). Test accommodations for students with disabilities: An analysis of the interaction hypothesis. *Review of Educational Research*, *75*(4), 457-490.

Schwartz (2013, December, 14). The selling of Attention Deficit Disorder. *New-York Times*. Retrieved on 26.5.2014

Timimi, S. (2008). Child psychiatry and its relationship with the pharmaceutical industry: theoretical and practical issues. *Advances in Psychiatric Treatment*, *14*(1), 3-9.

Thurlow, M. L., Lazarus, S. S., Thompson, S. J., & Morse, A. B. (2005). State policies on assessment participation and accommodations for students with disabilities. *The Journal of Special Education*, *38*(4), 232-240.

Tsal, Y., Shalev, L., & Mevorach, C. (2005). The Diversity of Attention Deficits in ADHD The Prevalence of Four Cognitive Factors in ADHD Versus Controls. *Journal of Learning Disabilities*, *38*(2), 142-157.

Wåhlstedt, C., Thorell, L. B., & Bohlin, G. (2009). Heterogeneity in ADHD: neuropsychological pathways, comorbidity and symptom domains. *Journal of Abnormal Child Psychology*, *37*(4), 551-564.

Zuriff, G. E. (2000). Extra examination time for students with learning disabilities: An examination of the maximum potential thesis. *Applied Measurement in Education*, *13*(1), 99-117.