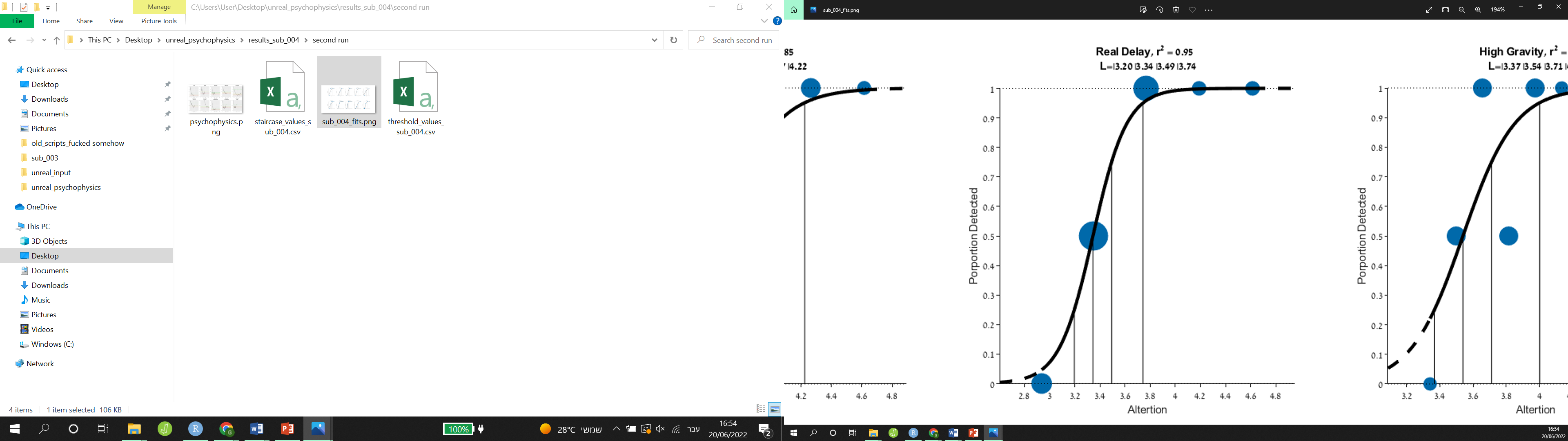
**סדר קבלת החלטות פסיכופיזיקה למריצות אנריל**

1. שיום קובץ תשובות הנבדק Answers.csv
2. להעביר את הקובץ לתיקיית נבדק בתוך unreal\_psychophysics
   1. דוגמה לכתובת: C:\Users\user\Desktop\unreal\_psychophysics\unreal\_04\sub\_003
3. מתוך unreal\_psychophysics לפתוח את הסקריפט R: run\_unreal\_psy.R
4. ללחוץ source
5. לכתוב ב-console: run\_unreal\_psy("C:\\Users\\User\\Desktop\\unreal\_psychophysics\\unreal\_04\\sub\_003")
6. ללחוץ Enter

הסקריפט אמור לייצר תיקיית תוצאות בתוך תיקיית הנבדק שמכילה:

1. Original\_threshold\_values\_sub\_003.csv
2. Staircase\_values\_sub\_003.csv
3. Threshold\_values\_adjusted\_sub\_003.csv
4. Sub\_003\_fits.png
5. Psychophysics\_sub\_003.png

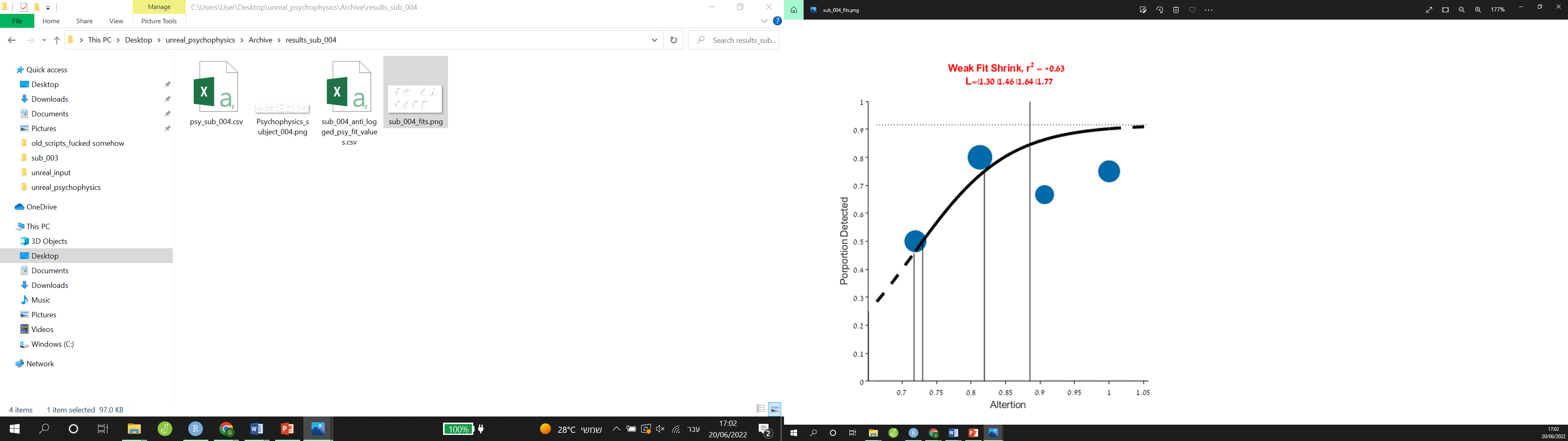
\*הקבצים בעזרתם מקבלים את ההחלטה האם צריך repeat הם קבצי 3-5. הגרפים המוצגים מטה הם מקובץ 4: הפיטים ממטלב.



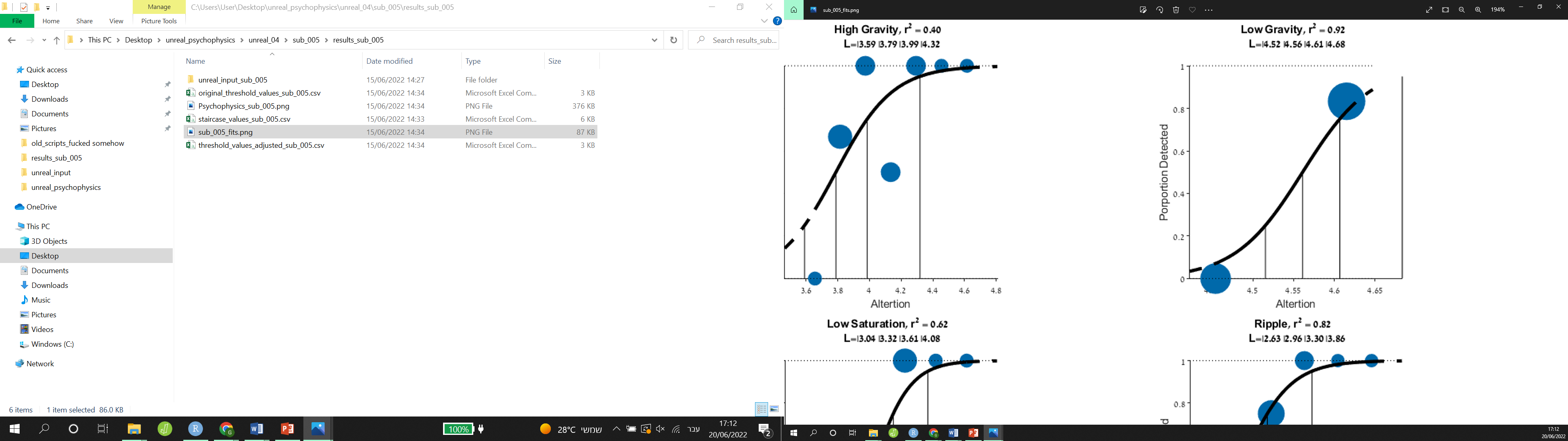
בגרף הראשון מעל מוצג פיט מוצלח במיוחד שעונה על כל הפרמטרים שלנו. המדדים שאכפת לנו מהם בגרפים של הפיטינג הם:

1. R² צריך להיות גדול מ-0.3, זה מדד שמייצג עד כמה העקומה המנובאת מתאימה לצורה של הדאטה שאספנו (העיגולים הכחולים).
2. אנחנו דורשים לפחות ערך אחד שנדגם בין האחוזון ה-85% של ציר ה-Y לאחוזון ה-15% (מסומן בעיגול אדום). כמו שרבים מאיתנו אולי למדו בקורס ניתוח אותו, הסיבה שאנו רוצים ערכים על השיפוע היא כדי שנהיה בטוחים שהעקומה המנובאת אמורה לעבור באזור הזה של הגרף. בלי ערכים על השיפוע, אפשר לצייר את העקומה בכל מנינ צורות ואין לנו ביטחון שהערכנו נכונה את רגישות הנבדק.
3. אנחנו רוצים צורת סיגמואיד עם שיפוע מתון ולא חד מדי (מכיוון שכך אנו מקבלים הפרדה טובה יותר בין ה-levels המחושבים לחלק ב' של הניסוי, שימו לב לחיצים הירוקים המייצגים מרחק סביר בין הרמות).

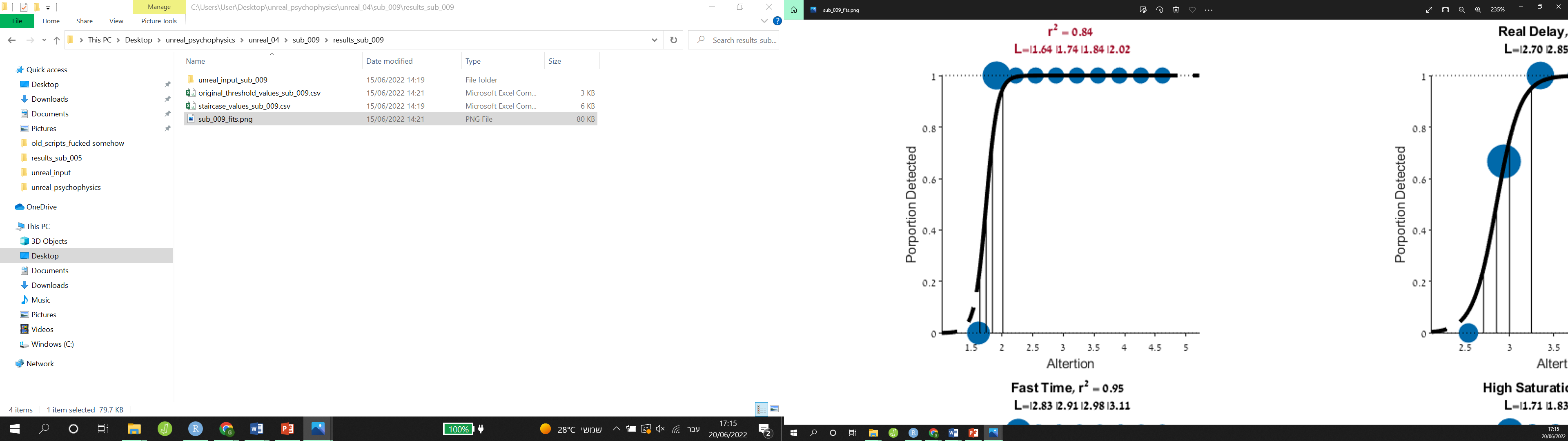
**סיבות לכך שפיט פסיכופיזי לא יהיה טוב**

1. הנבדק פספס את התנאי בתחילת הסטיירקייס ובעקבות כך הפיט מנבא שהנבדק לא יראה את הערכים החזקים של התנאי. שימו לב איך זה משפיע על ה- r² להיות שלילי.
2. לא נדגמו מספיק ערכים כדי לייצר פיט סביר, שימו לב שבמקרה הזה ה- r²נראה דווקא בסדר, אבל זה רק מכיוון שהערכים שנדגמו נופלים קרוב לעקומה המנובאת. בעצם זה פיט גרוע, כיוון שדגמנו רק שני ערכים ואין לנו ערך על השיפוע. המשמעות היא שאנחנו לא באמת אפיינו את הרגישות של הנבדק, אין לנו מושג מה באמת קורה בערכים באזור העיגול המקווקו השחור.

* בנוסף, מכיוון שהנבדק לא ראה חלק מהטריילים החזקים, הפיט מנבא לנו שערך ה-95% זיהוי הוא מחוץ לסקאלה של הניסוי. מסומן בעיגול המקווקו האדום. במקרה כזה, הסקריפט R שלנו יתקן את הthreshold העליון כך שהוא לא יהיה גדול יותר מהערך הראשון בסטיירקייס האחרון שהנבדק ביצע בתנאי הזה. בין אם הרצנו עליו סטיירקייס רגיל או שהרצנו עליו סטיירקייס מצעד אחד גבוה יותר.



1. הנבדק ראה או לא ראה את כל הערכים שנדגמו, לא הייתה שונות בתשובות, שימו לב שגם במקרה הזה ה- r²היה בסדר, כלומר העקומה מנבאת היטב את הדאטה. אבל אין לנו שיפוע מתון של הדגימות, ואין ערכים על השיפוע עצמו, כלומר לא מצאנו את הערכים של התנאי שהנבדק לפעמים רואה ולפעמים לא רואה. עוד תוצאה של המקרה הזה היא שה-levels המחושבים לחלק ב' של הניסוי מאוד צפופים, מה שלפעמים מעיד שהנבדק לא יבדיל בין הרמות בניסוי הסובייקטיבי וידרג את הכל אותו דבר בממוצע.

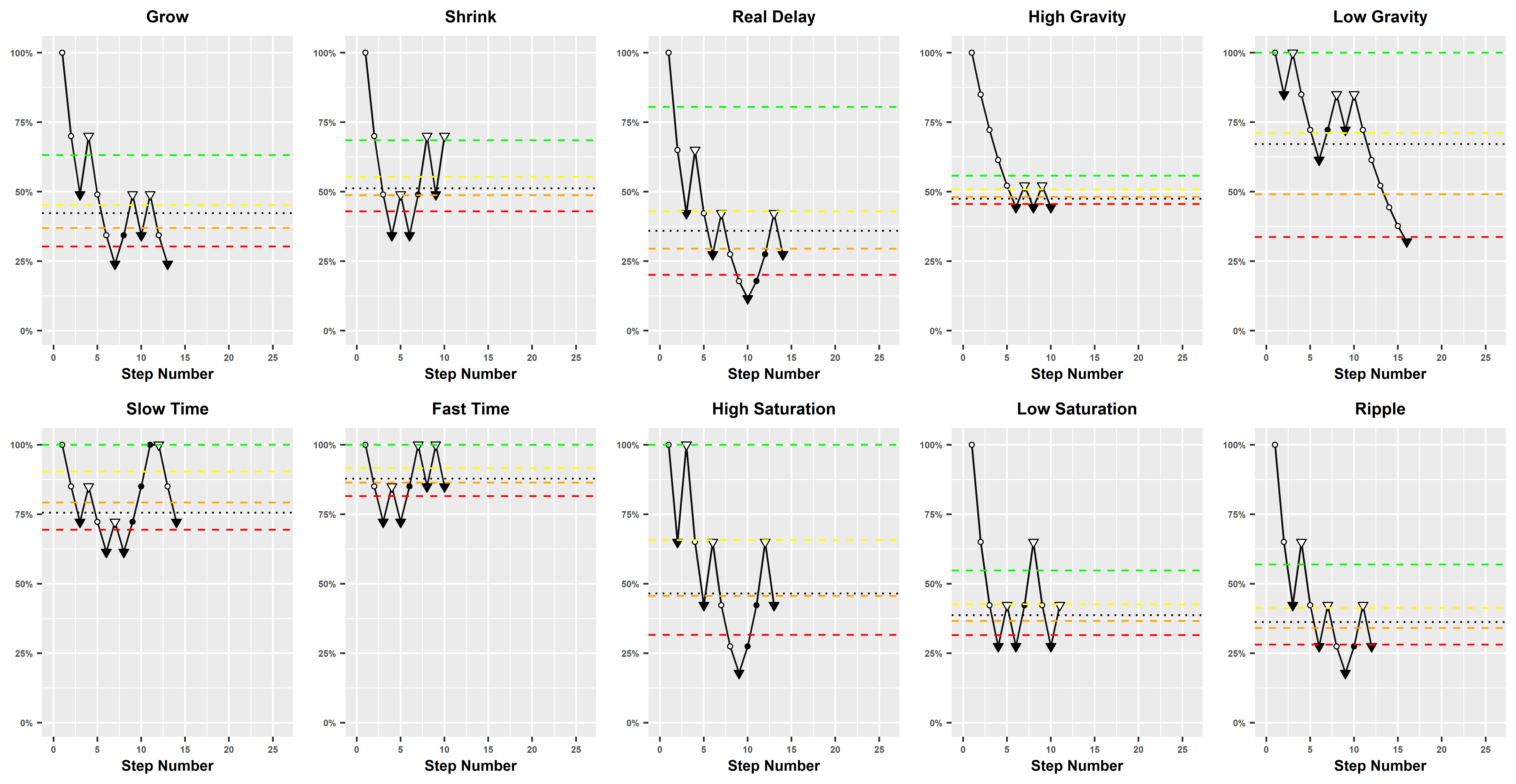


רק במקרה הראשון אנו יכולים להמשיך עם הניסוי כרגיל, במקרים 2 ו-3 נצטרך לעשות repeat לתנאי הלא מוצלח. כמה פעמים נעשה repeat? השיא שלנו עד היום הוא 4-5 הרצות חוזרות של אותו תנאי 😊 ההוראה הגורפת שלנו כרגע היא להריץ עד שיוצא כמו שצריך (כמו במקרה 1).

**הסתייגויות מהשיטה:** אפשר להסתכל על השיטה כאילו אנו 'מאמנים' נבדק חלש במקרים כאלה או שמדובר בצורה מסוימת של p-hacking. אבל צריך לזכור כמה דברים-

1. הניסוי שלנו כולו מבוסס על אפיון הרגישות של נבדקים בתנאים השונים על מנת שנוכל לשלוט בפרמטרים ולהשוות בין תנאים ובין נבדקים בצורה אובייקטיבית. איננו מחפשים מדגם שייתן תוצאה מובהקת, אלא מדגם שיאפשר להעריך נכונה את הרגישות הזאת.
2. במידה ולא הצלחנו לייצר פיט מוצלח לפי הקריטריונים של המחקר שאנחנו הצבנו, אנחנו מסתכנים בכך שהנבדק לא יחווה את הרמות השונות בתנאי בצורה מדורגת (יחווה במקום זאת הבדלים קיצוניים מדי או ללא הבדלים בכלל בין רמות).
3. אנחנו מסתכנים בביקורת מצד חוקרים אחרים על כך שלא עמדנו בקריטריונים של עצמנו.
4. אנחנו מסתכנים בביקורת על כך שלא אמדנו את הרגישות של הנבדק כמו שצריך.
5. מעל הכל, חשוב להבין שבניסוי פסיכופיזי אמיתי שבא לאפיין במדויק רגישות של בני אדם לגירוי מסוים דוגמים בסביבות 100 טריילים לסטיירקייס. כלומר, הפסיכופיזיקה שלנו קצרה בצורה קיצונית (בממוצע 15 טריילים לתנאי) ולכן יש הצדקה מחקרית לכך שפעמים רבות הפיט שלנו לא מוצלח (מעט מדי נתונים) ומצריך חזרה נוספת על מנת לדייק אותו.

במידה והחלטנו שהנבדק לא הצליח בתנאי מסוים, אנו עוברים להסתכל על גרף הסטיירקייס של התנאי הנ"ל:



הקו השחור המלא מייצג את המסלול שהנבדק עבר לאורך הסטיירקייס.

צעד לבן מייצג ערך שהנבדק זיהה בהצלחה, צעד שחור מייצג ערך שהנבדק לא זיהה. משולשים מייצגים נקודות רוורסיה.

הקווים הצבעוניים לרוחב הגרף מייצגים את הערך של הרמות שחושבו לתנאי מתוך הפיט שנעשה במטלב. הקו הירוק=95% זיהוי, הצהוב=75%, כתום=50% (JND), אדום=25%.

הקו השחור המקווקו מייצג את הערך הממוצע של נק' הרוורסיה.

שימו לב, קשה להגדיר מהו סטיירקייס מוצלח מבחינה גרפית. כעקרון אנו שואפים לגרף בירידה מתונה, עם דגימה מזגזגת כאשר מתחילים להתקרב לאזור ה-JND המיוחל. אבל הזיגזוג הזה (איפה שנקו' הרוורסיה מתחילות להופיע) יכול לקרות בכל מיני אזורים של הגרף, זה תלוי ברגישות הנבדק לתנאי המסוים.

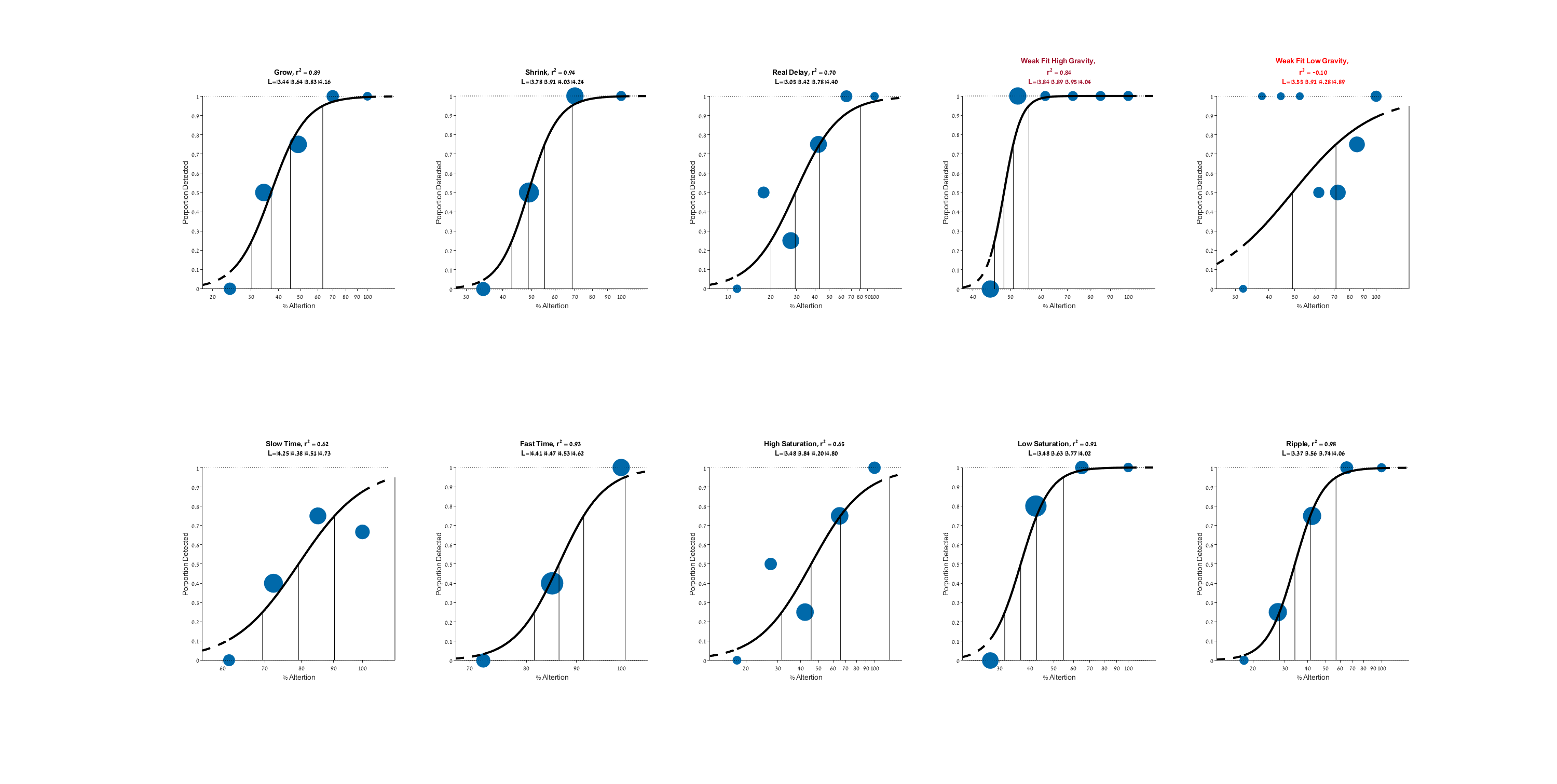
במיוחד פוגע בנו אם ישנם טריילים בתחילת הסטיירקייס (בצעד הראשון בדר"כ) בהם הנבדק דיווח שאינו מזהה את התנאי. לרוב אם מדובר בצעד יחיד, אנו מסתכלים על מקרים אלה כטעויות של קשב או הבנה, מכיוון שאנו דאגנו לבחור ערכי התחלה בכל תנאי שרוב האנשים מסוגלים לראות בקלות. אולם אם זה חוזר על עצמו בכמה טריילים חזקים, אנו מסיקים שהנבדק צריך סטיירקייס שמתחיל קצת גבוה יותר ובהתאם אם צריך repeat נבצע אותו עם סטיירקייס שמתחיל צעד אחד גבוה יותר.

אני איישר לכולנו קו כאן ואגדיר שאם הייתה יותר מטעות אחת בזיהוי הערך הגבוה ביותר בסטיירקייס, נבצע repeat בצעד גבוה יותר. שימו לב, ישנם שני קבצי staircasecommands בתיקיית ה-repeat בניסוי. כרגע אנו מגדירים במי מהם להשתמש באופן ידני על ידי שינוי השם לקובץ שאיננו רוצים למשהו שאינו staircasecommands.

\*למה לא לעשות repeat צעד גבוה יותר גרוף לכולם? מפני שהערך ההתחלתי של סטיירקייס משפיע על התוצאה הסופית, לכן אם לא נראה שיש צורך בכך נעדיף להמנע מזה.

שימו לב, מצורפים כאן למטה הפיטים של הנבדק שתואמים לגרפים של הסטיירקייסים למעלה. ניתן לראות שבגרביטציה גבוהה ונמוכה לא הוצאנו לו פיט מוצלח (כותרות אדומות/בורדו). בנוסף, אפשר לראות שב-4 תנאים (זמן מהיר/איטי, סטורציה גבוהה וגם עבור גרביטציה נמוכה) הערך של רמת ה-95% זיהוי מנובא מחוץ לסקאלה של הסטיירקייס (מעבר ל-100 על ציר האיקס, מסומן בעיגולים האדומים).

משום כך הם יעברו adjustment אוטומטי: בגרפים של הסטיירקייסים רואים שקו ה-95% (הירוק) הוצב בכוונה לא מעל ערך ה-100% כמו שמטלב הציע אלא בדיוק על ה-100% (מסומן בעיגולים כחולים). אנו עושים זאת על מנת שלא להציג לנבדקים ערכים גבוהים יותר ממה שהוצג להם בשלב הפסיכופיזיקה. כך אנו מנסים לשמר את הציפיות של הנבדק לניסוי הבא.



**אקסטרפולציה**

החלק המקווקו בגרפים של מטלב הוא אקסטרפולציה (מסומן בעיגולים הצהובים), כלומר לא נדגמו באמת ערכים בחלקים האלה של הגרף. זו היא תוספת מלאכותית של המשך העקומה שמטלב מנבא ואנו נמנעים מלהשתמש בערכים שנופלים בחלקים האלה של הגרף היות ואין לנו מידע שיאשר אם העקומה באמת נראית כך באזורים האלה.

**Threshold\_values\_adjusted**

קובץ המכיל בתוכו רשימה שמית מסודרת לפי הרמות של הערכים שהעברנו לצורך הניסוי השני כתוצאה מהפיטים. למען וידוא נוסף אחרון של התוצאות הפסיכופיזיות אפשר לבחון את הקובץ הזה בעיניים ולוודא שלא מופיעים בו ערכים חריגים במיוחד (למרות שהקוד כבר עושה זאת עבורנו, אבל תמיד יש הפתעות במחקר הזה).

דוגמאות למציאות שמצאנו בעבר הן:

* ערכים שליליים בסקאלה שאיננה שלילית
* ערכים מספריים במאות או אלפים שאין סיכוי שאמורים להופיע
* מספרים דימיוניים (יופיעו כמספר+מספר דימיוני, גם לא אמור לקרות יותר)
* אין מספר? NA
* אותו מספר בכמה רמות (לא אמור לקרות ואסור שיקרה)

\*ראיתם כמעט אותו מספר בכמה רמות וזה הדאיג אתכם? אל תדאגו זה בסדר גמור וקורה כל הזמן.

את תהליך קבלת ההחלטות בפסיכופיזיקה נתעד בדף נפרד שיצורף לדפי ההרצה שלנו.