Article: Doornaert, E. E., Mohamad, A. E. C., Johal, G., Allman, B. L., Möhrle, D., & Schmid, S. (2024). Not a Deficit, Just Different: Prepulse Inhibition Disruptions in Autism Depend on Startle Stimulus Intensities. *Eneuro*, *11*(9).

Test 1 – Nuisance variable analysis (Sex)

"There was no effect of sex or interaction effect between genotype and sex on the maximum startle response (sex, p= 0.1320, F(1,47)= 2.350; genotype × sex, p = 0.1262, F(1,47)= 2.424). Therefore, sex was collapsed to examine the effect of genotype on startle magnitude. Cntnap2 KO rats had a greater maximum startle response than their WT counterparts ($p < 0.0001$; Fig. 3A,C)." (Doornaert et al., 2024)

- Here, the authors interpret the non-significant p-value as evidence that sex has no effect. The p-value indicates insufficient evidence to detect an effect and should not be used to conclude that there is no effect at all. Additionally, the choice to collapse groups is presented as a logical consequence of this hypothesis test, rather than a pragmatic modeling decision which is framing their decision as a binary rule dependent on the p-value.
- A more revised statement of these results could be: "The effect of sex or interaction between genotype and sex was statistically unclear. Given the lack of clear statistical evidence, subsequent analyses were conducted with pooled data across sex to examine the effect of genotype on startle."

Test 2 – Main effects, interactions, post-hoc comparisons

"For the startle threshold, there was a significant main effect of genotype and sex, as well as an interaction effect between genotype and sex (genotype, $p < 0.0001$, $F(1,47) = 62.35$; sex, $p = 0.0001$, $F(1,47) = 17.88$; genotype × sex, $p = 0.0001$, $F(1,47) = 18.80$). Whereas Cntnap2 KO males showed a lower startle threshold than WT males, there was no difference between Cntnap2 WT and KO females (males, $p < 0.0001$; females, $p = 0.1283$; Fig. 3B,D)." (Doornaert et al., 2024)

- Here, no effect sizes and measures of uncertainty, such as confidence intervals, were presented. This makes it difficult to interpret the magnitude of the observed effects and to identify how variable the responses are. Additionally, the "no difference" they found within females is based entirely on the p-value which indicates insufficient evidence rather than a certainty of no effect.
- Revised statement: "The startle threshold differed by genotype and sex, with strong evidence indicating a genotype x sex interaction. To explore this, follow-up comparisons were done which found that Cntnap2 KO males exhibited a lower

startle threshold than WT males, while genotype differences in females remained statistically unclear."

Test 3 – 3-way Interaction (sex × genotype × startle stimulus)

"With the 75 dB prepulse, a significant interaction effect was found between sex and startle stimulus level, as well as between sex, genotype, and startle stimulus level (sex × startle stimulus, $p < 0.0001$, $F(5,235) = 13.70$; sex × genotype × startle stimulus, $p < 0.0001$, $F(5,235) = 6.230$)." (Doornaert et al., 2024)

- In this test, they discuss a three-way interaction (sex x genotype x startle stimulus) without the interpretation of the direction or pattern of the effect. Also, many F-values are presented here with no context which makes it hard to assess the results. For instance, without effect sizes or confidence intervals we do not know whether these effects are driven by large mean differences or variance.
- Revised: "With the 75 dB prepulse, a three-way ANOVA indicated an interaction between sex, genotype and startle stimulus level suggesting that genotype effects on startle magnitude varied across stimulus intensities and sex."

References

Doornaert, E. E., Mohamad, A. E.-C., Johal, G., Allman, B. L., Möhrle, D., & Schmid, S. (2024). Not a Deficit, Just Different: Prepulse Inhibition Disruptions in Autism Depend on Startle Stimulus Intensities. *eNeuro*, *11*(9), ENEURO.0179-24.2024. https://doi.org/10.1523/ENEURO.0179-24.2024