Cohen, J., Cohen, P., West, S. G., & **Aiken, L. S.** (2013). *Applied multiple regression/correlation analysis for the behavioural sciences*. Mahwah, NJ: Erlbaum. DOI: https://doi.org/10.4324/9780203774441

Cumming, G. (2005). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. New York, NY: Routledge.

David, F. N., & **Johnson, N. L.** (1951). The effect of non-normality on the power function of the f-test in the analysis of variance. *Biometrika*, *38*(1–2), 43–57. DOI: https://doi.org/10.2307/2332316

DeCarlo, L. T. (1997). On the meaning and use of kurtosis. *Psychological Methods*, *2*(3), 292–307. DOI: https://doi.org/10.1037//1082-989X.2.3.292

Delacre, M., Lakens, D., & **Leys, C.** (2017). Why psychologists should by default use Welch's t-test instead of student's t-test. *International Review of Social Psychology*, *30*(1), 92–101. DOI: https://doi.org/10.5334/irsp.82

Erceg-Hurn, D. M., & **Mirosevich, V. M.** (2008). Modern robust statistical methods: An easy way to maximize the accuracy and power of your research. *American Psychologist*, *63*(7), 591–601. DOI: https://doi.org/10.1037/0003-066X.63.7.591

Glass, G. V., Peckham, P. D., & **Sanders, J. R.** (1972). Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance. *Review of Educational Research*, *42*(3), 237–288. DOI: https://doi.org/10.3102/00346543042003237

Green, E. G., Deschamps, J.-C., & **Páez, D.** (2005). Variation of individualism and collectivism within and between 20 countries: A typological analysis. *Journal of Cross-Cultural Psychology*, *36*(3), 321–339. DOI: https://doi.org/10.1177/0022022104273654

Grey, S., & **Mathews, A.** (2000). Effects of training on interpretation of emotional ambiguity. *The Quarterly Journal of Experimental Psychology*, *53*(4), 1143–1162. DOI: https://doi.org/10.1080/713755937

Grissom, R. (2000). Heterogeneity of variance in clinical data. *Journal of Consulting and Clinical Psychology*, *68*(1), 155–165. DOI: https://doi.org/10.1037//0022-006X.68.1.155

Haar, J. M., Russo, M., Suñe, A., & **Ollier-Malaterre, A.** (2014). Outcomes of work-life balance on job satisfaction, life satisfaction and mental health: A study across seven cultures. *Journal of Vocational Behavior*, *85*(3), 361–373. DOI: https://doi.org/10.1016/j.jvb.2014.08.010

Harwell, M. R., Rubinstein, E. N., Hayes, W. S., & **Olds, C. C.** (1992). Summarizing Monte Carlo results in methodological research: The one- and two-factor fixed effects anova cases. *Journal of Educational Statistics*, *17*(4), 315–339. DOI: https://doi.org/10.3102/10769986017004315

Henrich, J., Heine, S. J., & **Norenzayan, A.** (2010). Most people are not weird. *Nature*, *466*, 29–29. DOI: https://doi.org/10.1038/466029a

Heun, R., Burkart, M., Maier, W., & **Bech, P.** (1999). Internal and external validity of the who well-being scale in the elderly general population. *Acta Psychiatrica Scandinavica*, *99*(3), 171–178. DOI: https://doi.org/10.1111/j.1600-0447.1999.tb00973.x

Hoekstra, R., Kiers, H. A., & **Johnson, A.** (2012). Are assumptions of well-known statistical techniques checked, and why (not)? *Frontiers in Psychology*, *3*(137), 1–9. DOI: https://doi.org/10.3389/fpsyg.2012.00137

Hsu, T.-C., & **Feldt, L. S.** (1969). The effect of limitations on the number of criterion score values on the significance level of the f-test. *American Educational Research Journal*, *6*(4), 515–527. DOI: https://doi.org/10.3102/00028312006004515

Keppel, G., & **Wickens, T. D.** (2004). *Design and analysis: A researcher's handbook*. Upper Saddle River, New Jersey: Prentice Hall.

Keselman, H., Huberty, C. J., Lix, L. M., Olejnik, S., Cribbie, R. A., Donahue, B., Levin, J. R., et al. (1998). Statistical practices of educational researchers: An analysis of their anova, manova, and ancova analy. *Review of Educational Research*, *68*(3), 350–386. DOI: https://doi.org/10.3102/00346543068003350

Koeser, S., & **Sczesny, S**. (2014). Promoting gender-fair language: The impact of arguments on language use, attitudes, and cognitions. *Journal of Language and Social Psychology*, *33*(5), 548–560. DOI: https://doi.org/10.1177/0261927X14541280

Liu, H. (2015). *Comparing welch anova, a kruskal-wallis test, and traditional anova in case of heterogeneity of variance* (PhD thesis). Virginia Commonwealth University.

Lix, L. M., Keselman, J. C., & **Keselman, H.** (1996). Consequences of assumption violations revisited: A quantitative review of alternatives to the one-way analysis of variance *f* test. *Review of Educational Research*, *66*(4), 579–619. DOI: https://doi.org/10.3102/00346543066004579

Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin*, *105*(1), 156–166. DOI: https://doi.org/10.1037/0033-2909.105.1.156

Montoya, D. Y., & **Briggs, E.** (2013). Shared ethnicity effects on service encounters: A study across three us subcultures. *Journal of Business Research*, *66*(3), 314–320. DOI: https://doi.org/10.1016/j.jbusres.2011.08.011

Nimon, K. F. (2012). Statistical assumptions of substantive analyses across the general linear model: A mini-review. *Frontiers in Psychology*, *3*(322), 1–5. DOI: https://doi.org/10.3389/fpsyg.2012.00322

Overall, J. E., Atlas, R. S., & **Gibson, J. M.** (1995). Tests that are robust against variance heterogeneity in k × 2 designs with unequal cell frequencies. *Psychological Reports*, *76*(3), 1011–1017. DOI: https://doi.org/10.2466/pr0.1995.76.3.1011

Quensel, C.-E. (1947). The validity of the *z*-criterion when the variates are taken from different normal populations. *Scandinavian Actuarial Journal*, *30*(1), 44–55. DOI: https://doi.org/10.1080/03461238.1947.10419648