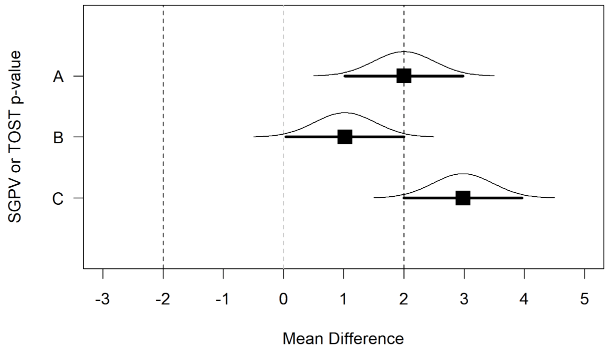

Figure 3. Means, normal distribution, and 95% CI for three example datasets that illustrate the relationship between *p*-values from TOST and SGPV.

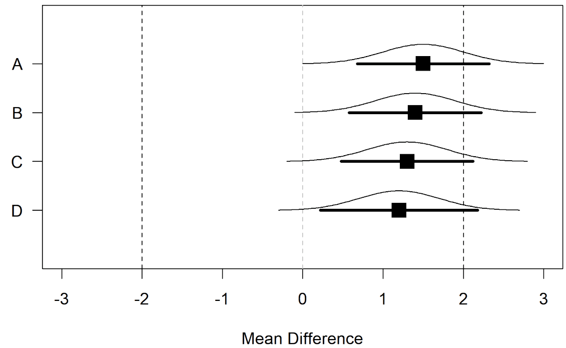

Figure 4. Means, normal distribution, and 95% CI for samples where the observed population mean is 1.5, 1.4, 1.3, and 1.2.

SGPV is 50% when the observed mean falls exactly on the lower or upper equivalence bound, because 50% of the symmetrical confidence interval overlaps with the equivalence range. When the observed mean equals the equivalence bound, the difference between the mean in the data and the equivalence bound is 0, the *t*-value for the equivalence test is also 0, and thus the *p*-value is 0.5 (situation A, Figure 3).

Two other points always have to overlap. When the 95% CI falls completely inside the equivalence region, and one endpoint of the confidence interval is exactly equal to one of the equivalence bounds (see situation B in Figure 3) the TOST *p*-value (which relies on a one-sided test) is always 0.025, and the SGPV is 1. Note that when sample sizes are small or equivalence bounds are narrow, small *p*-values for the TOST or a SGPV = 1 might not be observed in practice if too few observations are collected. The third point where the SGPV and the *p*-value from the TOST procedure should overlap is where the 95% CI falls completely outside of the equivalence range, but one endpoint of the confidence interval is equal to the equivalence bound (see situation C in Figure 3), when the *p*-value will always be 0.975, and the SGPV is 0. Note that this situation is in essence a minimum-effect test (Murphy, Myors, & Wolach, 2014). The goal of a minimum-effect is not just to reject a difference of zero, but to reject the smallest effect size of interest (i.e., the equivalence bounds). An equivalence test and minimum effect test against the same equivalence bound are complementary, and when a TOST *p*-value is larger than 0.975, the *p*-value for the minimum effect test is smaller than 0.05 (and therefore the minimum effect test provides no additional information that can not be derived from the *p*-value from the equivalence test). The SGPV summarizes the information from an equivalence test (and the complementary minimum-effect test). These can be two relevant questions to ask, although it often makes sense to combine an equivalence test and a null-hypothesis test instead (Lakens et al., 2018).

For example, in Figure 4 we have plotted four SGPV's. From A to D the SGPV is 0.76, 0.81, 0.86, and 0.91. The difference in the percentage of overlap between A and B (-0.05) is identical to the difference in the percentage of overlap between C and D as the mean gets 0.1 closer to the test value (-0.05). As the observed mean in a one-sample *t*-test lies closer to the test value, from situation A to D, the difference in the overlap changes uniformly. As we move the observed mean closer to the test value in steps of 0.1 across A to D the *p*-value calculated for normally distributed data are not uniformly distributed. The probability of observing data more extreme than the upper bound of 2 is (from A to D) 0.16, 0.12, 0.08, and 0.05. As we can see, the difference between A and B (0.04) is not the same as the difference between C and D (0.03). Indeed, the difference in *p*-values is the largest as you start at $p = 0.5$ (when the observed mean falls on the test value), which is why the line in Figure 1 is the steepest at $p = 0.5$. Note that where the SGPV reaches 1 or 0, *p*-values closely approximate 0 and 1, but never reach these values.

When different *p*-values for equivalence tests yield the same SGPV

There are three situations where *p*-values for TOST differentiate between observed results, while the SGPV does not differentiate. The first two situations were discussed before and can be seen in Figure 1. When