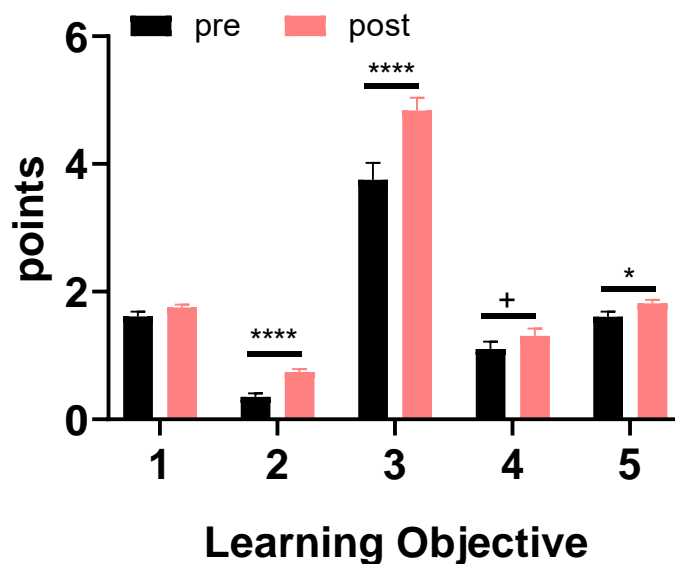


**Supplementary Figure 1:** Pre and Post test scores with all questions included. Students had higher test scores in the post-test ( $mdn = 10.8$ ) compared to the pre-test ( $mdn = 8.8$ ) and this difference was statistically significant via a Wilcoxin matched-pairs signed rank test ( $z = -4.9$ ,  $p < 0.0001$ ,  $r = 0.69$ ). Error bars are reported as standard error of the mean.



**Supplementary Figure 2:** Pre and Post test scores on each learning objective, with Question 13 omitted (part of LO4). Question 2 is included since it is the only question that assesses LO2. Students had greater scores for LO2 ( $z = -4.6$ ,  $p < 0.0001$ ,  $r = 0.81$ ), LO3 ( $z = -3.8$ ,  $p < 0.0001$ ,  $r = 0.60$ ), and 5 ( $z = -2.3$ ,  $p < 0.05$ ,  $r = 0.46$ ) during the post-test compared to the pre-test. LO4 showed a trending increase ( $z = -1.66$ ,  $p < 0.05$ ,  $r = 0.32$ ) while there was not an increase in performance in LO1 ( $z = -1.19$ ,  $p = 0.23$ ,  $r = 0.21$ ). Error bars are reported as standard error of the mean.