1. (ii) and (iii)

2. (iii)

3. (a) Since 1400 is less than two standard deviations above the mean 1050 (noticeably less than two standard deviations, so it has a Z-score lower than 1.96), it is within the nearest 95% of observations in proximity to the mean. That means it does not reach as high as the 98^{th} percentile.

   (b) The two standardized Z-scores:

$$\text{You:} \quad \frac{1300-1050}{200} = 1.25$$

$$\text{Your friend:} \quad \frac{25-20}{6} = 0.833$$

   Your performance is the one with the higher Z-score, and should be ranked as higher.

4. (ii)

5. (a) The left tail consists of observations as extreme or more extreme than $\widehat{p}_1 - \widehat{p}_2 = -0.246$, and these occur with relative frequency 0.028. For a two-tailed alternative, we double this, arriving at a P-value of 0.056.

   (b) One approach is to write "Intelligence" on 29 slips of paper and "effort" on 30 slips, placing all 59 in one bag and mixing them up. In advance, you can have decided the first 15 draws (doing so without replacement) from the bag will be treated as the 15 liars, and the rest as the ones who do not lie. Once the bag is emptied, you can calculate the proportion of "Intelligence" slips considered honest (i.e., what proportion of the slips marked that way were among the final 44 slips drawn), calling that $\widehat{p}_1$, and calculating the analogous proportion of slips marked "Effort". Finally, you subtract the two: $\widehat{p}_1 - \widehat{p}_2$.

   (c) We first compute the pooled proportion:

$$\tilde{p} = \frac{44}{59},$$

   and using that along with sample sizes n_1, n_2, we get an estimated standard error

$$\text{SE}_{\widehat{p}_1-\widehat{p}_2} = \sqrt{\frac{44}{59}\cdot\frac{15}{59}\cdot\left(\frac{1}{29}+\frac{1}{30}\right)} \doteq 0.1134,$$

   not a huge amount different than the SE reported on the randomization distribution dot plot, but it is a bit different. To standardize:

$$z = \frac{-0.246}{0.1134} = -2.1693.$$

   (d) In this case,

$$n_1\widehat{p}_1 = 18, \quad n_1(1-\widehat{p}_1) = 11, \quad n_2\widehat{p}_2 = 26, \quad n_2(1-\widehat{p}_2) = 4,$$

   so one of the rules of thumb for justifying a normal model is not met. So, the randomization method is somewhat more reliable, as it just uses frequencies and makes no assumptions.

6. (a) (iv)

   (b) Neither sample size reaches 30, the level at which we are encouraged to *assume* normality. If we knew the underlying populations were normal, then we could rest at ease, but that is beyond the given information. There just is not enough to go on.

   (c) We would use

```r
qt(0.98, df=17)
```

(d) Since the null value 0 is inside the 96% confidence interval and not in the rejection region, our P-value must be greater than 0.04.

7. (iv)

8. (a) A bootstrap distribution for a sample mean $\overline{x}$ repeatedly draws a bootstrap sample, and computes $\overline{x}$ as the bootstrap statistic. Doing so makes the original sample mean, not the null value, the center of this distribution. The important extra step is to add to each $\overline{x}$ a number equal to the difference between the null value 5.7 and the test statistic, given in part (b) to be 5.552.

   (b) The standardized test statistic is

$$\frac{5.552 - 5.7}{0.552/\sqrt{50}} \doteq -1.8959.$$

   (c) *(3 pts)* Write an RStudio command that generates a P-value, using your test statistic from part (b). We could use

```r
pt(-1.8959, df=49) * 2
```