

1、自变量：文字内容和打印颜色是否一致

因变量：测试者说出列表中覆盖文字的颜色时间

2、 $H_0: \mu_1 = \mu_2$ ——覆盖文字的颜色和文字字面意义是否相符，并不会对说出颜色的总体平均时间产生影响。

$H_A: \mu_1 < \mu_2$ ——当覆盖文字的颜色和文字字面意义不一致时说出颜色的总体平均时间要比覆盖文字的颜色和文字字面意义一致时说出颜色的总体平均时间要长。

其中， $\mu_1$  表示覆盖文字的颜色和文字字面意义一致时，说出颜色的总体平均时间；

$\mu_2$  表示覆盖文字的颜色和文字字面意义不一致时，说出颜色的总体平均时间。

根据测试，测试本身并没有得到总体的测试结果，只有样本测试的结果，无法得知总体偏差，且样本容量小于 30，所以应该选择 t 检验；同一名测试者同时进行了两种条件不同的测试，属于相依样本中的 **Two conditions** 的样本类型；由于假设检验中的对立假设是：当覆盖文字的颜色和文字字面意义不一致时说出颜色的总体平均时间要比覆盖文字的颜色和文字字面意义一致时说出颜色的总体平均时间要长，所以应该选择正向单尾检验。

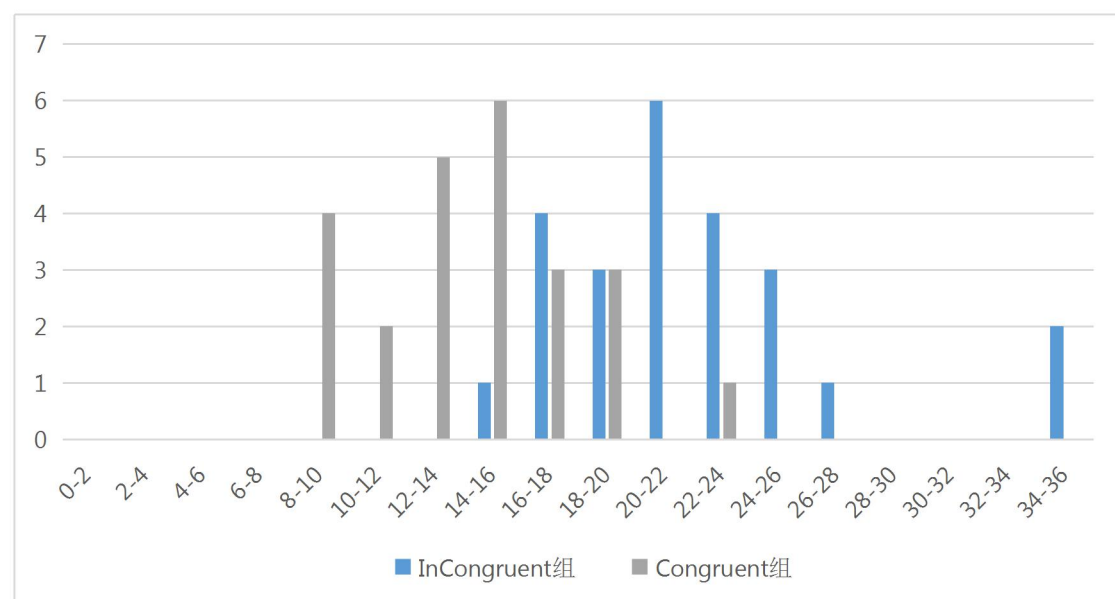
首先，样本是随机抽取的，而且只有一个时间变量，其属于连续变量。其次，测试者之间相互独立，而且测试本身只是得到每一位测试者在文字内容和打印颜色在一致和不一致两种情况下说出打印颜色的时间，测试者之间相互独立。最后由测试的性质及样本的分布直方图推测总体分布直方图应近似正态分布，且尚无足够理由推翻这一推测。由上边三点保证 t 检验中相依样本的可行性和结果的可信度。

3、对于 Congruent 组：其均数约为 14.05，中位数约为 14.36，其数据集的标准差约为 3.56，IQR 为 4.31。

对于 Incongruent 组：其均数约为 22.02，中位数约为 21.02，其数据集的标准差为 4.80，IQR 为 5.33。

4、

数据可视化：



由图可以看出 Congruent 组的数据分布偏向于正态分布，且大部分数据集中在 14-16 的区间内。

由图可见，Incongruent 组的数据分布偏向于正态分布，且数据主要集中在 20-22 的区间内，并存在两个疑似异常值落在 34-36 的区间内。

5、

在  $\alpha = 0.05$  的正向单尾检验的情况下，t 检验的临界值为 1.714。

$\therefore \bar{x}_I = 22.02$  (Incongruent 组的均值)

$\bar{x}_C = 14.05$  (Congruent 组的均值)

$\therefore$  根据样本均值差异假设总体均值差异  $\mu_I - \mu_C = 7.97$

根据  $D = x_{I_i} - x_{C_i}$ ，求得  $\bar{D} = 7.96$

$$\therefore S = \sqrt{\frac{\sum (D_i - \bar{D})^2}{n-1}} = \sqrt{\frac{544.33}{23}} = 4.86482691$$

$$\therefore t = \frac{\mu_I - \mu_C}{S / \sqrt{n}} = \frac{7.97}{4.86482691 / \sqrt{24}} = 8.02$$

$$\therefore 8.02 > 1.714$$

$$\therefore P < 0.05, \text{ 且 } P \approx 0$$

同时求得 95%CI =  $(-\infty, 9.70)$

由于  $8.02 > 1.714$ ， $P < 0.05$ ，所以拒绝零假设。

结论：在显著水平 0.05 下，覆盖文字的颜色和文字字面意义不一致时说出颜色的总体平均时间要比覆盖文字的颜色和文字字面意义一致时说出颜色的总体平均时间要长，且两种条件下说出颜色的总体平均时间之差以 95% 的置信度落在置信区间  $(-\infty, 9.70)$  内。

结果与期望一致。

6、

原因：人们通常会注意到所看到的信息较为显眼、奇异的部分，而大脑也会先处理先接收到的信息按照习惯思路经过习惯的信息加工然后进行反应，如例中的文字的字面意义；而当要求对接收到的信息需要进行不同以往的处理时，如增加信息转换的判断条件等会在一定程度上增加大脑处理这段信息的时间，导致从做出判断到处理信息最后到反应这一过程时间增加，如例中制造出覆盖颜色和文字字面意义不一样这样的情况，要测试者正确说出覆盖颜色的名称。而且，人们在处理得到的信息的时候，有时会收到同时接受到的别的信息的干扰，导致筛选有用信息的时间增加，从而使处理信息前这一段时间有所增加，增长总体反应时间，如例中字面意思通常表示一种颜色以干扰覆盖颜色。

类似的测试：如双语 stroop 范式测试：利用两种语言的色词，要求被试用母语和第二语言分别对两种语言的色词进行颜色命名。