1. What's the difference between a parametric statistic (e.g., ANOVA, Pearsons' product-moment correlation coefficient) and a nonparametric statistic (e.g., Spearman's rho)? [1 pt]

**A nonparametric statistic makes fewer assumptions about the nature of the data. Thus, ordinal or nominal data would be interpretable with certain nonparametric statistics. Parametric statistics, however, may presume that the data are interval or ordinal.**

2. Place, et al. (2008) wrote an article that appeared in *Psychological Science* titled: "The ability to judge the romantic interest of others." The abstract read:

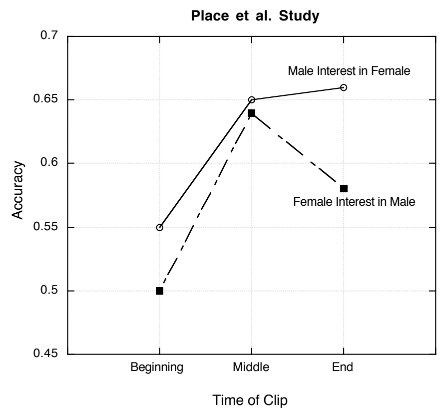
The ability to judge another individual's romantic interest level—both toward oneself and toward others—is an adaptively important skill when choosing a suitable mate to pursue. We tested this ability using videos of individuals on speed dates as stimuli. Male and female observers were equally good at predicting interest levels, but they were more accurate when predicting male interest: Predictions of female interest were just above chance. Observers predicted interest successfully using stimuli as short as 10 s, and they performed best when watching clips of the middle or end of the speed date. There was considerable variability between daters, with some being very easy to read and others apparently masking their true intentions. Variability between observers was also found. The results suggest that the ability to read nonverbal behavior quickly in mate choice is present not only for individuals in the interaction, but also for third-party observers.

To simplify the study and place it in a format that is consistent with your knowledge, let's imagine the study as a 2x3 independent groups design. Forty-eight participants (24 male and 24 female) viewed a short video clip of a speed dating couple. They were shown only a portion of the speed dating interaction from the beginning, middle, or end of the interaction. Half the participants rated only the males in the interaction (Male's interest in female) and the other half of the participants rated only the females in the interaction (Female's interest in male). The dependent variable is accuracy in correctly judging the romantic interest of the target person in the video (on a scale of 0 = totally inaccurate to 1 = totally accurate). Below is a source table that analyzes data similar to those found by Place et al. Complete the source table and interpret the results as completely as you can. [15 pts]

**Descriptive Statistics**

Dependent Variable:Accuracy

| Gender | Time | Mean | Std. Deviation | N |
|---|---|---|---|---|
| Male Interest in Female | Beginning of clip | .5500 | .04000 | 8 |
| | Middle of clip | .6513 | .01959 | 8 |
| | End of clip | .6600 | .03464 | 8 |
| | Total | .6204 | .05974 | 24 |
| Female Interest in Male | Beginning of clip | .4975 | .02375 | 8 |
| | Middle of clip | .6437 | .02722 | 8 |
| | End of clip | .5812 | .02949 | 8 |
| | Total | .5742 | .06639 | 24 |
| Total | Beginning of clip | .5238 | .04177 | 16 |
| | Middle of clip | .6475 | .02324 | 16 |
| | End of clip | .6206 | .05118 | 16 |
| | Total | .5973 | .06671 | 48 |




## Tests of Between-Subjects Effects

Dependent Variable:Accuracy

| Source | Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared | Noncent. Parameter | Observed Power[b] |
|---|---|---|---|---|---|---|---|---|
| Gender | .026 | 1 | .026 | 26 | .000 | .406 | 28.739 | .999 |
| Time | .136 | 2 | .068 | 68 | .000 | .783 | 151.798 | 1.000 |
| Gender * Time | .010 | 2 | .005 | 5 | .006 | .217 | 11.630 | .846 |
| Error | .042 | 42 | .001 | | | | | |
| Corrected Total | .214 | 47 | | | | | | |

**Homogeneity of Variance: Could either compute Hartley's $F_{Max}$ (4.21) and compare to $F_{Max\ Crit}$ (10.8) and determine that there was no reason to be concerned about heterogeneity of variance (and use $\alpha = .05$) OR note that all the *F*s were significant with *p*s less than .01, so they would be significant even in the face of heterogeneity of variance.**

$$HSD = 4.2\sqrt{\frac{.001}{8}} = .047$$

**There was a significant main effect of Gender, $F(1,42) = 26$, $MSE = .001$, $p < .001$, $\eta^2 = .406$. There was also a significant main effect of Time, $F(2,42) = 68$, $p < .001$, $\eta^2 = .783$. Finally, there was a significant interaction of Gender and Time, $F(2,42) = 5$, $p = .006$, $\eta^2 = .217$. Post hoc tests using Tukey's HSD revealed that at the beginning and end of the clip, males were significantly more accurate ($M = .550$ and $M = .660$, respectively) than females ($M = .498$ and $M = .581$, respectively). However, at the middle of the clip, males ($M = .651$) and females ($M = .644$) did not differ significantly.**

3. In an article by Shackman et al. (2009) in *Psychological Science* entitled “Right dorsolateral prefrontal cortical activity and behavioral intervention,” the abstract reads:

Individuals show marked variation in their responses to threat. Such individual differences in behavioral inhibition play a profound role in mental and physical well-being. Behavioral inhibition is thought to reflect variation in the sensitivity of a distributed neural system responsible for generating anxiety and organizing defensive responses to threat and punishment. Although progress has been made in identifying the key constituents of this behavioral inhibition system in humans, the involvement of dorsolateral prefrontal cortex (DLPFC) remains unclear. Here, we acquired self-reported Behavioral Inhibition System Sensitivity scores and high-resolution electroencephalography from a large sample. Using the enhanced spatial resolution afforded by source modeling techniques, we show that individuals with greater tonic (resting) activity in right-posterior DLPFC rate themselves as more behaviorally inhibited. This observation provides novel support for recent conceptualizations of behavioral inhibition and clues to the mechanisms that might underlie variation in threat-induced negative affect.

From the PASW analysis below (of data that mimic those found in the Shackman et al. study), interpret the results as completely as you can. How many participants are in the study? If a person received a score of 20 on the Behavioral Inhibition Scale (BIS), what would you predict that person’s EEG asymmetry score to be? If a person received a score of 12 on the Behavioral Inhibition Scale (BIS), what would you predict that person’s EEG asymmetry score to be? What is the function of the standard error of estimate? [10 pts]

**Model Summary**

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
|---|---|---|---|---|
| 1 | .448[a] | .201 | .184 | .13669 |

a. Predictors: (Constant), Behavioral Inhibition Scale

**ANOVA[b]**

| Model | | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| 1 | Regression | .226 | 1 | .226 | 12.083 | .001[a] |
| | Residual | .897 | 48 | .019 | | |
| | Total | 1.123 | 49 | | | |

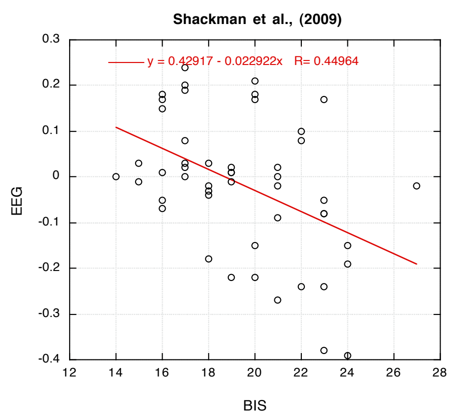
a. Predictors: (Constant), Behavioral Inhibition Scale

b. Dependent Variable: EEG asymmetry

**Coefficients[a]**

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
|---|---|---|---|---|---|---|
| | | B | Std. Error | Beta | | |
| 1 | (Constant) | .427 | .129 | | 3.319 | .002 |
| | Behavioral Inhibition Scale | -.023 | .007 | -.448 | -3.476 | .001 |

a. Dependent Variable: EEG asymmetry




There was a significant negative linear relationship between EEG and BIS, *r*(48) = -.448, *p* = .001. The sample size, *n*, was 50. The coefficient of determination, $r^2$, was .201. The standard error of estimate, *SEE*, was .137. The *SEE* allows one to determine the accuracy of predictions, with a smaller *SEE* indicating that the points are closer to the line of best fit, so the accuracy would be higher.

Using the regression equation (Y = -.023X + .427), one would predict that a BIS score of 20 would yield an EEG score of -.033. Because you didn’t observe a BIS score of 12, you could either decline to make a prediction, because of that fact, or you could make a prediction of .151, but only if the trend continued.

4. [G&W] Is there a relationship between cognitive ability and social rank? It may be the case for birds. Boogert, Reader, and Laland (2006) measured social rank and individual learning ability for a group of starlings. The following data represent results similar to those obtained in the study. Analyze the data as completely as possible. [15 pts]

| Starling | Social Rank | Learning Score | Ranked Social | Ranked Learning | SxL |
|---|---|---|---|---|---|
| A | 1 | 3 | 1 | 1.5 | 1.5 |
| B | 3 | 10 | 5.5 | 5 | 27.5 |
| C | 2 | 7 | 3 | 4 | 12 |
| D | 3 | 11 | 5.5 | 6 | 33 |
| E | 5 | 19 | 9.5 | 10 | 95 |
| F | 4 | 17 | 7.5 | 8.5 | 63.75 |
| G | 5 | 17 | 9.5 | 8.5 | 80.75 |
| H | 2 | 4 | 3 | 3 | 9 |
| I | 4 | 12 | 7.5 | 7 | 52.5 |
| J | 2 | 3 | 3 | 1.5 | 4.5 |
| $\Sigma X$ | | | 55 | 55 | 379.5 |
| $\Sigma X^2$ | | | 381.5 | 384 | |
| *SS* | | | 79 | 81.5 | |

$$rho = \frac{379.5 - \frac{(55)(55)}{10}}{\sqrt{\left(381.5 - \frac{55^2}{10}\right)\left(384 - \frac{55^2}{10}\right)}} = \frac{77}{\sqrt{(79)(81.5)}} = .96$$

**$r_{s\ Crit}$ (10) = .648**

**Because (the absolute value of) Spearman's rho is greater than the critical value of rho, you would reject $H_0$ and conclude that there is a significant positive relationship. Starlings with the higher social rank (higher numbers indicate greater social rank) also had higher learning scores.**

5. Some questions related to repeated measures designs:

a. In a single factor repeated measures design with 7 levels of the factor (A), how many participants would you need to conduct the study if you wanted a minimum of 15 scores per condition (cell)? [2 pts]

**With incomplete counterbalancing there would be 14 orders, so you'd need 28 people (using each order twice).**

b. Suppose that you conduct the study (as in *a* above). Complete the resulting source table below. [5 pts]

| **Source** | ***SS*** | ***df*** | ***MS*** | ***F*** |
|---|---|---|---|---|
| Between (A) | 120 | 6 | 20 | 10 |
| Within | 567 | 189 | | |
| Subject | 243 | 27 | | |
| Error | 324 | 162 | 2 | |
| Total | 687 | 195 | | |

c. Use the above source table to illustrate why it is that the repeated measures analysis will typically be more powerful than the independent groups analysis. That is, what component of the source table is largely responsible for making the repeated measures analysis more powerful? (Show why.) [2 pts]

**The Subject effect (removing individual differences)…so, if $SS_{Subj}$ is small, then $SS_{Error}$ will be larger and $F$ will be smaller. If $SS_{Subj}$ is large, then $SS_{Error}$ will be smaller and $F$ will be larger.**

d. What component of the above source table is most similar to the interaction term in a two-factor ANOVA? [1 pt]

**The Error term, which is the interaction between Treatment and Subject.**

e. Why is it that you typically cannot conduct a repeated measures study for a characteristic of a person (e.g., gender, IQ)? [1 pt]

**Because you cannot (typically) manipulate people's characteristics, you can't them as factors in repeated measures designs.**

f. Under what other circumstances would you not be able to conduct an experiment as a repeated measures design? [1 pt]

**When your manipulation produces a permanent change, then you cannot use that IV in a repeated measures design. For example, comparing the removal of one area of a frog's brain with the removal of another area of the frog's brain on some behavior.**

g. In an independent groups ANOVA, the best estimate of population variance ($\sigma^2$) is $MS_{Within}$ ($MS_{Error}$). Why is it that $MS_{Error}$ in a repeated measures ANOVA is *not* a good estimate of population variance? [2 pts]

**Population variability is due to individual differences and random variability. However, $MS_{Error}$ in a repeated measures design is due to random variability, thus it will typically underestimate the population variance.**

6. Master et al. (2009) reported in *Psychological Science* a study entitled "A picture's worth: Partner photographs reduce experimentally induced pain." Their introduction follows (no abstract)…

Social support is associated with reduced pain experience across several domains (Cogan & Spinnato, 1988; Kulik & Mahler, 1989; Zaza & Baine, 2002); intriguingly, a handful of experimental studies suggest that this connection may reflect a causal relationship. Participants who received interactive support during a cold pressor task reported less pain than participants who completed the task alone or engaged in nonsupportive interactions (Brown, Sheffield, Leary, & Robinson, 2003; Jackson, Iezzi, Chen, Ebnet, & Eglitis, 2005). Moreover, the mere presence of another supportive individual (vs. being alone) reduced pain ratings in a cold pressor task (Brown et al., 2003; but see McClelland & McCubbin, 2008) and reduced pain ratings among fibromyalgia patients following stimulation to a painful body site (Montoya, Larbig, Braun, Preissl, & Birbaumer, 2004).

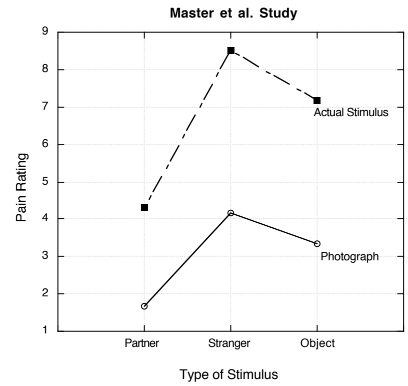
Could the same pain-attenuating effects of social support be observed by merely activating the mental representation of a supportive other? Previous work has shown that activating mental representations of important others can produce effects similar to those created by the actual presence of these individuals (Fitzsimons & Bargh, 2003; Mikulincer & Shaver, 2001). Building on this research, the current study examined whether simply viewing a photograph of one's romantic partner could reduce physical-pain experience. We examined how this condition compared with one that is more consistent with previous conceptualizations of social support—one in which the participant held her partner's hand.

I'll modify the design of the study slightly, to make it consistent with your statistical knowledge. Women who were in relationships sat in front of a screen that prevented them from seeing their left arm, which was behind the screen. Thermal pain was delivered to the arm and rated by the participants (DV = pain rating, with lower numbers indicating lower pain ratings). In some cases, the pain was rated while either: a. holding her partner's hand behind the curtain, b. holding a stranger's hand (the experimenter) behind the curtain, or c. holding an object (squeezable ball) behind the curtain. In other cases, the pain was rated while either: a. looking at a photograph of her partner, b. looking at a photograph of a male stranger (matched for ethnicity of actual partner), or c. looking at a photograph of an object. Thus, you can think of this study as a 2x3 independent groups design (Focal Stimulus: Real vs. Photograph and Type of Stimulus: Partner, Stranger, Inanimate Object). Complete the source table below and interpret the data as completely as you can. [15 pts]

**Descriptive Statistics**

Dependent Variable:Pain

| Focal Stimulus | Type of Stimulus | Mean | Std. Deviation | N |
|---|---|---|---|---|
| Photograph of Stimulus | Partner | 1.6667 | .81650 | 6 |
| | Stranger | 4.1667 | .75277 | 6 |
| | Inanimate Object | 3.3333 | 1.03280 | 6 |
| | Total | 3.0556 | 1.34917 | 18 |
| Actual Stimulus | Partner | 4.3333 | 1.03280 | 6 |
| | Stranger | 8.5000 | 1.04881 | 6 |
| | Inanimate Object | 7.1667 | .75277 | 6 |
| | Total | 6.6667 | 2.00000 | 18 |
| Total | Partner | 3.0000 | 1.65145 | 12 |
| | Stranger | 6.3333 | 2.42462 | 12 |
| | Inanimate Object | 5.2500 | 2.17945 | 12 |
| | Total | 4.8611 | 2.48599 | 36 |




## Tests of Between-Subjects Effects

Dependent Variable:Pain

| Source | Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared | Noncent. Parameter | Observed Power[b] |
|---|---|---|---|---|---|---|---|---|
| FocalStim | 117.4 | 1 | 117.4 | 139.9 | .000 | .823 | 139.901 | 1.000 |
| TypeStim | 69.4 | 2 | 34.7 | 41.36 | .000 | .734 | 82.715 | 1.000 |
| FocalStim * TypeStim | 4.4 | 2 | 2.2 | 2.62 | .090 | .148 | 5.232 | .481 |
| Error | 25.2 | 30 | .839 | | | | | |
| Corrected Total | 216.4 | 35 | | | | | | |

**Homogeneity of Variance: Either compute $F_{Max}$ (1.94) and compare to $F_{Max\ Crit}$ (18.7), which would lead you to use a = .05 OR determine that the two significant effects (main effects) were significant with *p* < .01.**

**For Type of Stimulus:** $HSD = 3.486\sqrt{\frac{.839}{12}} = .922$

**There was a significant main effect for Focal Stimulus, F(1,30) = 139.9, MSE = .839, p < .001, $\eta^2$ = .823. The pain rating for the photograph of the stimulus was significantly lower (*M* = 3.056) than for the actual stimulus (*M* = 6.667). There was also a main effect for the Type of Stimulus, F(2,30) = 41.36, p < .001, $\eta^2$ = .734. Post hoc tests using Tukey's HSD indicate that the pain ratings were significantly higher when the stimulus was a stranger (*M* = 6.333) than an inanimate object (*M* = 5.25) or the partner (*M* = 3.000). The pain ratings were higher when the person focused on the inanimate object than the partner. There was no significant interaction, *F*(2,30) = 2.62, *p* = .09, $\eta^2$ = .148.**