1. Farzin, et al. (2010) wrote an article that appeared in *Psychological Science* titled: "Spatial resolution of conscious visual perception in infants." The abstract read:

Humans' conscious awareness of objects in their visual periphery is limited. This limit is not entirely the result of reduced visual acuity. Rather, it is primarily caused by crowding—the difficulty identifying an object when it is surrounded by clutter. The effect of crowding on visual awareness in infants has yet to be explored. Do infants, for example, have a fine-grained "spotlight," as adults do, or do infants have a diffuse "lantern" that sets limits on what they can register in their visual periphery? We designed an eye-tracking paradigm to psychophysically measure crowding in infants between 6 months and 15 months of age. We showed infants pairs of faces at three eccentricities, in the presence or absence of flankers, and recorded infants' first saccade from central fixation to either face. Infants could discriminate faces in the periphery, and flankers impaired this ability. We found that the effective spatial resolution of infants' visual perception increased with age, but was only half that of adults.

Fig. 2. Mooney face stimuli shown to participants in Experiment 1. All faces used were cropped to fit ellipses, as shown in (a). The top row in (a) includes original faces used in Mooney's 1957 study. On each trial, one upright face and one inverted face were presented. In the uncrowded condition (b), these two faces were displayed by themselves, and in the crowded condition (c), six flanker images surrounded each face. Eccentricity of the faces was varied; in the examples shown here, the faces are at 3° eccentricity.

For our purposes, the two conditions were Age (6-, 9-, 12-, and 15-month olds) and Crowding (b = uncrowded and c = crowded). To simplify one of their studies and place it in a format that is consistent with your knowledge, let's imagine the study as a 2x4 independent groups design. Half of the two hundred and forty participants (60 6-month-olds, 60 9-month-olds, 60 12-month-olds, and 60 15-month-olds) viewed the Mooney faces as either Uncrowded or Crowded (so 30 6-month-olds saw Uncrowded and 30 saw Crowded, etc.). The dependent variable is the mean threshold of eccentricity (departure from center, so greater eccentricity is more peripheral) for detecting the face as upright or upside down. Thus, higher eccentricity for thresholds indicates that the infants could detect the upright face (for example) at a greater eccentricity (more peripheral presentation). Complete the source table below and interpret the results as completely as you can [15 pts]

Descriptive Statistics

Dependent Variable:Threshold

Age	Crowding	Mean	Std. Deviation	N
6 mo	Uncrowded	1.0000	.30057	30
	Crowded	1.8000	.21972	30
	Total	1.4000	.48047	60
9 mo	Uncrowded	3.1440	.63951	30
	Crowded	1.0000	.30057	30
	Total	2.0720	1.18916	60
12 mo	Uncrowded	3.0000	.43470	30
	Crowded	1.0500	.31045	30
	Total	2.0250	1.05214	60
15 mo	Uncrowded	3.3033	.37553	30
	Crowded	2.5700	.42520	30
	Total	2.9367	.54305	60
Total	Uncrowded	2.6118	1.04259	120
	Crowded	1.6050	.71806	120
	Total	2.1084	1.02588	240

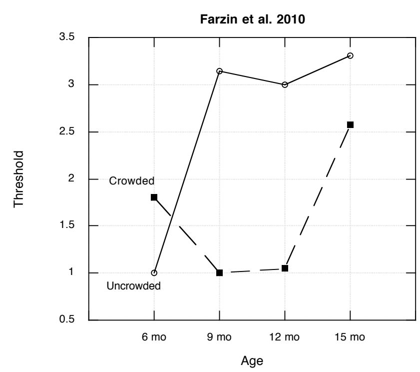
Levene's Test of Equality of Error Variances[a]

Dependent Variable:Threshold

F	df1	df2	Sig.
2.487	7	232	.018

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + Age + Crowding + Age * Crowding



Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Observed Power[b]
Age	71.768	3	23.9	153.6	.000	.665	1.000
Crowding	60.823	1	60.8	390.7	.000	.627	1.000
Age * Crowding	82.832	3	27.6	177.4	.000	.696	1.000
Error	36.108	232	.16				
Corrected Total	251.531	239					

Homogeneity of Variance: The Levene Test is significant, so you'd be concerned about heterogeneity of variance and proceed with α = .01. However, all the effects are significant with $p < .001$, so they'd be significant regardless of the α-level you choose. [Had you chosen to compute Hartley's F_{Max}, you'd have arrived at the same conclusion, with $F_{Max} = 8.2$ and $F_{Max\ Crit} = 3.12$.]

$$HSD = 4.35\sqrt{\frac{.16}{30}} = .31$$

There is a significant main effect of Age, $F(3,232) = 153.6$, $MSE = .16$, $p < .001$, $\eta^2 = .665$. There is a significant main effect of Crowding, $F(1,232) = 390.7$, $p < .001$, $\eta^2 = .627$. There was also a significant interaction between Age and Crowding, $F(3,232) = 177.4$, $p < .001$, $\eta^2 = .696$. Using Tukey's HSD, the interaction (as seen in the figure) arises because at 6 months of age, the eccentricity threshold for Crowded pictures is greater than for Uncrowded pictures. However, for 9-, 12-, and 15-month-olds, the eccentricity threshold for Uncrowded pictures is greater than for Crowded pictures. OR For Crowded pictures, the eccentricity threshold was greater for 15-month-olds than for all other ages, 6-month-olds had higher eccentricity thresholds than 9- and 12-month-olds. However, for the Uncrowded pictures, 9-, 12-, and 15-month-olds had higher eccentricity thresholds than did 6-month-olds.

2. In an article by Roelofs et al. (2010) in *Psychological Science* entitled "Facing freeze: Social threat induces bodily freeze in humans," the abstract reads:

Freezing is a common defensive response in animals threatened by predators. It is characterized by reduced body motion and decreased heart rate (bradycardia). However, despite the relevance of animal defense models in human stress research, studies have not shown whether social threat cues elicit similar freeze-like responses in humans. We investigated body sway and heart rate in 50 female participants while they were standing on a stabilometric force platform and viewing cues that were socially threatening, socially neutral, and socially affiliative (angry, neutral, and happy faces, respectively). Posturographic analyses showed that angry faces (compared with neutral faces and happy faces) induced significant reductions in body sway. In addition, the reduced body sway for angry faces was accompanied by bradycardia and correlated significantly with subjective anxiety. Together, these findings indicate that spontaneous body responses to social threat cues involve freeze-like behavior in humans that mimics animal freeze responses. These findings open avenues for studying human freeze responses in relation to various sociobiological markers and social-affective disorders.

The female subjects completed the State-Trait Anxiety Index (STAI), which is a 20-item questionnaire with each response on a 4-pt scale. Thus, as used in this study, the scores could range from 0 (low anxiety) to 80 (high anxiety). The subjects looked at faces while their body sway was measured. The faces exhibited a neutral, happy, or angry emotion. The body-sway difference score in the analysis below is calculated by subtracting body-sway variability in the neutral faces block from the same measure in the angry-faces block. Thus, a difference score of 0 would indicate that body sway was the same when viewing neutral and angry faces. However, a negative difference score would indicate that body sway was reduced for angry faces compared to neutral faces. And a positive difference score would indicate that body sway was greater for angry faces compared to neutral faces.

From the SPSS analysis below (of data that mimic those found in the Roelofs 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 30 on the STAI, what would you predict that person's Body-Sway Difference score to be? If a person received a score of 70 on the STAI, what would you predict that person's Body-Sway Difference 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	.345[a]	.119	.101	.43757

a. Predictors: (Constant), Anxiety

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1.240	1	1.240	6.475	.014[a]
	Residual	9.190	48	.191		
	Total	10.430	49			

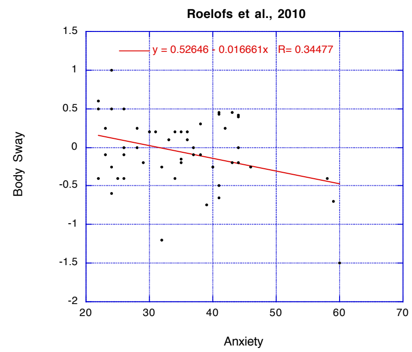
a. Predictors: (Constant), Anxiety

b. Dependent Variable: BodySway

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	.526	.237		2.224	.031
	Anxiety	-.017	.007	-.345	-2.545	.014

a. Dependent Variable: BodySway



There was a significant negative linear relationship between body sway and anxiety, $r(48) = -.345$, $p = .014$. The coefficient of determination was $r^2 = .119$. The $n = 50$. The regression equation is:

$$\hat{y} = (-.017)(x) + .526$$

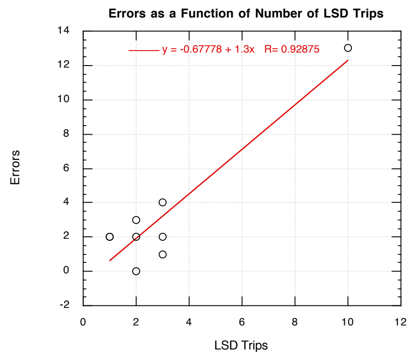
Thus, for an anxiety level of 30, body sway would be predicted to be .016. For anxiety level of 70, which is beyond the observed levels of anxiety, you could predict body sway of -.66 if the trend continues. Or, you could simply decline to predict beyond values you've observed.

The function of the standard error of estimate is to gauge the accuracy of the predictions made by your regression equation.

3. Dr. C. D. Snow suspects that LSD affects the speech center in the brain. Specifically, she believes that repeated use of LSD reduces a person's ability to retrieve verbal information. To see if she can obtain any evidence regarding the relationship between LSD and retrieval of verbal information, Dr. Snow advertises for people who have taken LSD at least once. Nine people of comparable IQ and education are selected from the applicants. All participants are given a 50-item test. Each item consists of a definition of a low-frequency English word; the participant's task is to produce the target word.

Here are the relevant data. Analyze these data as completely as you can. [15 pts]

Subject	Number of reported LSD trips	Number of failures to produce the target word (errors)	XY
1	2	2	4
2	3	2	6
3	1	2	2
4	2	0	0
5	10	13	130
6	1	2	2
7	3	1	3
8	3	4	12
9	2	3	6
Sum	27	29	165
SS	60	117.56	



$$r = \frac{165 - \frac{(27)(29)}{9}}{\sqrt{(60)(117.56)}} = .929$$

With $r_{Crit}(7) = .666$, we would reject H_0: $\rho = 0$ and conclude that there is a significant positive linear relationship (with $r_{Obtained} \geq r_{Critical}$). The regression equation would be:

$$\hat{y} = 1.3x - .7$$

$$SEE = \sqrt{\frac{(.137)(117.56)}{7}} = 1.517$$

Taking the time to plot the data should lead you to question the one extreme point (outlier?). Without that point, the linear relationship would disappear ($r = .14$). Moreover, you should think of the design problem of looking only at people who have taken LSD. For appropriate comparison, it would make sense to include people who had never taken LSD because their data may look just like those who had taken LSD (but you wouldn't know if you didn't include their data).

4. Dr. Mai Ayes was interested in studying the effects of task difficulty and sleep deprivation on performance, using a completely between (independent groups) design. The amounts of sleep deprivation that she decided to use are: 36, 48, 60, 72, and 84 hours. That is, participants were awake without sleep for one of those periods before being tested on either an easy, a moderate, or a difficult task. She measured performance on a 9-point scale (1 = lousy performance <-> 9 = excellent performance). Analyze these data as completely as you can. [15 pts]

Descriptive Statistics

Dependent Variable:Performance

HrsDep	Difficulty	Mean	Std. Deviation	N
36 Hrs	Easy	6.3000	1.70294	10
	Moderate	5.4000	.96609	10
	Difficult	2.9000	1.66333	10
	Total	4.8667	2.04658	30
48 Hrs	Easy	6.1000	.99443	10
	Moderate	4.3000	.94868	10
	Difficult	2.8000	1.47573	10
	Total	4.4000	1.77337	30
60 Hrs	Easy	4.9000	.87560	10
	Moderate	3.5000	1.08012	10
	Difficult	2.0000	.94281	10
	Total	3.4667	1.52527	30
72 Hrs	Easy	3.6000	.84327	10
	Moderate	2.7000	.67495	10
	Difficult	2.1000	.87560	10
	Total	2.8000	.99655	30
84 Hrs	Easy	3.6000	1.34990	10
	Moderate	2.4000	.51640	10
	Difficult	1.8000	.78881	10
	Total	2.6000	1.19193	30
Total	Easy	4.9000	1.64441	50
	Moderate	3.6600	1.37929	50
	Difficult	2.3200	1.23619	50
	Total	3.6267	1.77055	150

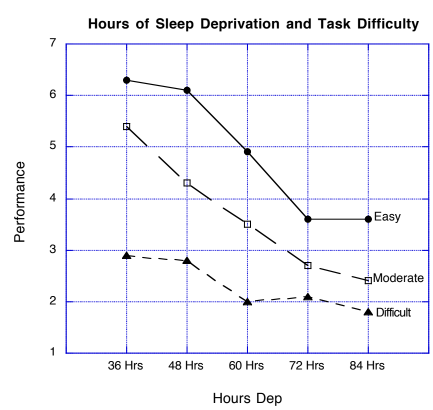
Levene's Test of Equality of Error Variances[a]

Dependent Variable:Performance

F	df1	df2	Sig.
2.406	14	135	.005

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + HrsDep + Difficulty + HrsDep * Difficulty



Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Observed Power[b]
HrsDep	116.97	4	29.24	24.187	.000	.417	1.000
Difficulty	166.51	2	83.25	68.862	.000	.505	1.000
HrsDep * Difficulty	20.45	8	2.56	2.114	.039	.111	.829
Error	163.22	135	1.209				
Corrected Total	467.15	149					

Homogeneity of variance: The Levene test is significant (*p* = .005), so you'd be concerned about heterogeneity of variance and proceed with α = .01. You'd reach the same conclusion had you chosen to compute Hartley's F_{Max} (F_{Max} = 10.9, $F_{Max\ Crit}$ = 10.7).

Hours Deprivation:	Difficulty:
$HSD = 3.92\sqrt{\frac{1.209}{30}} = .787$	$HSD = 3.35\sqrt{\frac{1.209}{50}} = .52$

There was a significant main effect for Hours of Deprivation, *F*(4,135) = 24.187, *MSE* = 1.209, *p* < .001, η^2 = .417. Post hoc tests using Tukey's HSD indicate that those deprived of sleep for 36 (*M* = 4.867) and 48 hours (*M* = 4.4) had significantly better performance than those who had been deprived of sleep for 60 (*M* = 3.467), 72 (*M* = 2.8), and 84 hours (*M* = 2.6). Those deprived of sleep for 60 hours performed significantly better than those deprived of sleep for 80 hours.

There was also a significant main effect for Task Difficulty, *F*(2,135) = 68.862, *p* < .001, η^2 = .505. Post hoc tests using Tukey's HSD indicate that performance was significantly better in the Easy task (*M* = 4.9) than on the Moderate (*M* = 3.66) or Difficult tasks (*M* = 2.32). Performance was significantly better on the Moderate task than on the Difficult task.

There was no significant interaction between Hours of Deprivation and Task Difficulty, *F*(8,135) = 2.56, *p* = .039, η^2 = .111.

You should also note that there should probably have been a control group that wasn't deprived of sleep at all.

5. Dr. John Goff is interested in whether or not the nature of a video produces differences in yawning. He first keeps his 11 participants awake for 36 hours straight. Dr. Goff then has the participants first watch a 30-minute Interesting video and then a 30-minute Uninteresting video. As they watch the videos, he observes the number of times that the participants yawn. Below are the data. Analyze the data as completely as you can. [20 pts.]

Subject	Interesting Video	Uninteresting Video	P
1	5	7	12
2	2	1	3
3	4	7	11
4	3	8	11
5	0	2	2
6	4	5	9
7	7	6	13
8	6	9	15
9	3	3	6
10	1	4	5
11	8	9	17
Sum	43	61	104
SS	60.9	76.7	

Source	***SS***	***df***	***MS***	***F***
Between	14.76	1	14.76	8.58
Within	137.6	20		
Subject	120.4	10		
Error	17.2	10	1.72	
Total	152.4	21		

With $F_{Crit}(1,10) = 4.97$ the decision would be to reject H_0: $\mu_{Interesting} = \mu_{Uninteresting}$.

There was a significant effect of the nature of the video on number of yawns, $F(1,10) = 8.58$, $MSE = 1.72$, $p < .05$, $\eta^2 = .46$. Those who saw the uninteresting video yawned significantly more often ($M = 5.55$) than those who saw the interesting video ($M = 3.91$).

To properly counterbalance this study would involve an equal number of participants. With $n = 11$, there couldn't have been proper counterbalancing.

6. Some questions related to repeated measures designs:

a. In a single factor repeated measures design with 8 levels of the factor (A), how many participants would you need to conduct the study if you wanted a minimum of 20 scores per condition (cell)? [2 pts]

24

b. Suppose that you conduct the study (as in *a* above). Complete the resulting source table below. [3 pts]

Source	***SS***	***df***	***MS***	***F***
Between (A)	140	7	20	10
Within	552	184		
Subject	230	23		
Error	322	161	2	
Total	692	191		

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]

SS_{Subjects} is removed from SS_{Within} to yield the SS_{Error}. Had SS_{Subjects} been larger (e.g., 391), then MS_{Error} would become 1 and F would increase to 20. However, had SS_{Subjects} been smaller (e.g., 69), then MS_{Error} would become 3 and F would decrease to 6.67.

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

SS_{Error} is the interaction between Treatment (Between) and Subject, so it's the most similar to $SS_{\text{Interaction}}$.

e. In an independent groups ANOVA, the best estimate of population variance (σ^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]

Because the MS_{Error} in a repeated measures ANOVA reflects only random variability, it wouldn't be a good estimate of σ^2, which is influenced by both individual differences and random variability.

7. Some random questions: [5 pts]
a. What is the name commonly used for the standard deviation of the sampling distribution of the mean?

Standard error

b. In a sample of $n = 25$ scores, you obtain a mean (M) of 5.3 and a standard deviation (s) of 5. Suppose that you want to test H_0: $\mu = 7$. Compute the appropriate statistic and test H_0.

$$s_{\bar{X}} = {}^{5}/_{\sqrt{25}} = 1$$
$$t = \frac{5.2 - 7}{1} = -1.7$$

With $t_{\text{Critical}}(24) = 2.06$, you'd retain H_0: $\mu = 7$.

c. If you want to detect a small effect, what would you need to have and how would you get it?

You'd need power, which you could get in a number of ways, including increasing *n*.