

Exam 1

PS 306, Spring 2003

1. Define an operational definition and then give an example of a specific operational definition from the Mook article or the video from the first day of class. If you needed to design a study investigating stress, what would be your first step in operationally defining stress? [5 pts]

According to your text, an operational definition "forms a crucial link between our ideas and the world" by defining crucial terms in the hypothesis in concrete terms. Following Kerlinger, Ray suggests two types of operational definition: measured and experimental. Measured operational definitions "may specify both *how* observations are to be made and *what* is to be observed and measured." Experimental operational definitions "describes how experimental procedures are to be followed."

Any clear example of an operational definition from Mook or from the video would be fine. The first step in operationally defining stress would be to read the literature to determine how researchers typically define stress operationally. It's always safest to work from an extant operational definition.

2. Dr. Will Parr was interested in the impact of stress on performance on a motor task. To that end, he randomly assigned people to one of three levels of stress (Low, Moderate, High) based on the level of shock threatened at the end of the experiment. People in the Low stress condition were told that they would receive a very mild shock if the number of holes-in-one they achieved was less than a secret number, which would be revealed at the end of the course. People in the Moderate stress condition were given the same instructions, but were told that the shock would be moderately painful, but, of course, no permanent tissue damage. People in the High stress condition were given the same instructions, but were told that the shock would be very painful. The DV used was the number of holes-in-one achieved on the challenging putt-putt golf course. Complete the analysis seen in the source table below, interpret the results as completely as you can, then tell Dr. Parr what he should do next. [10 pts]

ANOVA Table for Performance

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Stress	2	.600	.300	1.227	.3089	2.455	.236
Residual	27	6.600	.244				

Means Table for Performance

Effect: Stress

	Count	Mean	Std. Dev.	Std. Err.
High	10	1.400	.516	.163
Low	10	1.700	.483	.153
Moderate	10	1.700	.483	.153

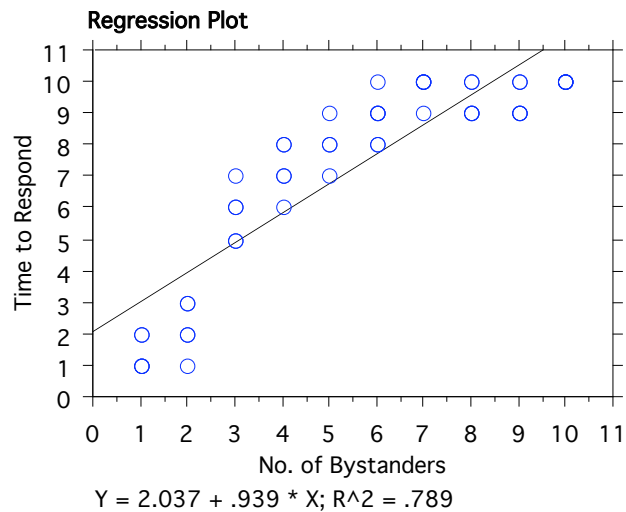
There is no significant effect of stress in this independent groups design ($F(2,27) = 1.23, p = .31$). That means that you had too little power in your experiment. First, you should note the absence of a No Stress control group. Inclusion of such a group would likely have increased the size of the treatment effect. You could also increase the treatment effect by making the High Stress group even more stressful (maybe by giving actual shocks of a high

intensity). Second, you would always look to an increased sample size as a way to increase power. In this case, with $n = 10$, increasing sample size would make a lot of sense. Finally, you should always look to ways that you might reduce MS_{Error} . Could you make your instructions clearer? Would it be worthwhile to make your groups more homogeneous?

3. We talked about the Darley and Latané study of bystander apathy as an ANOVA. However, you would also be able to conduct a study that looked at many different levels of number of bystander with a correlation/regression analysis. That is, suppose that you randomly assigned $n = 5$ participants to each of 10 levels of number of “bystanders” in the communication study used by Darley & Latané (1, 2, 3, 4, 5, 6, 7, 8, 9, and 10). (So the 1 bystander condition would be the participant and one other person, who was your confederate.) As your dependent variable you then measured the number of minutes until a participant went to investigate the choking confederate. The number of minutes to respond for each of the participants was entered into the analysis seen below. (Because some participants responded in the same amount of time, some of the circles overlap in the scattergram below.) Interpret the results as completely and clearly as you can. If a person were in this situation with 3 other people, how quickly would that person respond? Would you feel comfortable making causal claims based on this study? [10 pts]

Regression Summary
Time to Respond vs. No. of Bystanders

Count	50
Num. Missing	0
R	.889
R Squared	.789
Adjusted R Squared	.785
RMS Residual	1.421



ANOVA Table

Time to Respond vs. No. of Bystanders

	DF	Sum of Squares	Mean Square	F-Value	P-Value
Regression	1	363.620	363.620	180.009	<.0001
Residual	48	96.960	2.020		
Total	49	460.580			

Regression Coefficients

Time to Respond vs. No. of Bystanders

	Coefficient	Std. Error	Std. Coeff.	t-Value	P-Value
Intercept	2.037	.435	2.037	4.678	<.0001
No. of Bystanders	.939	.070	.889	13.417	<.0001

There is a significant positive linear relationship between Number of Bystanders and Time to Respond, with longer reaction times associated with greater numbers of bystanders, $r(48) = .89, p < .05$. Using the regression equation ($Y = 2.037 + .939X$), if there were 3 bystanders present, I would predict that the time to respond would be 4.9 minutes. Because of the nature of the design, you would be able to make causal claims in this study. It's important to understand that it's not the analysis that determines the ability to make causal claims, but the nature of the design. So a correlational analysis is appropriate for this study, but it's not a correlational design...but an actual experimental design.

4. You've read the Mook article, with its extended discussion of external validity. First, define external validity and then describe the source of external invalidity in the Brown & Hanlon study (acquisition of grammatical speech through parental feedback) and the Hecht study (dark adaptation). Then, tell me why Mook (and you?) think that these specific examples of external invalidity are not of great concern. [10 pts]

Clear explication of the concept of external validity and the reason that concerns about the lack of external validity may not be crucial are central to the answer to this question.

5. Briefly, but clearly, indicate why a nonmanipulated characteristic of a participant, such as intelligence, may be of interest to a psychologist, but that it doesn't lend itself to studies that can result in causal claims about the impact of intelligence on some dependent variable. [5 pts]

Psychologists are interested in many nonmanipulated characteristics of participants (e.g., age, gender, race) because they are often readily measured and are likely to play some role in explaining human behavior. Nonetheless, they are problematic because they are not (and cannot be) manipulated. Why is that a problem? Think about intelligence, for example. Imagine two people...one with high intelligence and one with low intelligence. What are the odds that these two people differ only in intelligence level? Quite low, eh? For example, you might imagine that the person with high intelligence would be more likely to have obtained a better education, may have experienced more success in dealing with life's difficulties, etc. Thus, if you observed a significant difference between people with low and high intelligence in terms of their scores on some dependent variable, you would not know if the difference was due to intelligence or to some other variable that is related to intelligence.

6. What is meant by a manipulation check? Under which experimental conditions would a manipulation check be very useful? Using a study from the Mook article, indicate how a manipulation check might have been used to buttress the study. [5 pts]

A manipulation check is used to determine the effectiveness of a manipulation when one cannot directly manipulate the IV. For example, if your IV is whether or not a person is wearing glasses (Argyle study), you would not need a manipulation check (it's obvious if the person is wearing glasses or not). However, if your IV was anxiety level (Higgins & Marlatt), you can't know for sure that your means of manipulation (threat of electric shock) actually led to greater or lesser amounts of anxiety. In the H&M study, they could have asked participants a number of questions at the end of the study, one of which might have been, "How anxious were you about the possibility that you would be shocked in the course of this study?" You could use a 9-pt rating scale for the participant's response.

7a. Suppose that you were interested in the impact of two different drugs (Drug X and Drug Y) on maze performance in rats. You decide to use a repeated measures design to yield the greatest power and efficiency in your study. Describe your study as clearly and completely as you can (number of conditions, number of rats, procedure you would use, etc.). [15 pts]

It's unlikely that you'd actually do this study as a repeated measures design. That said, it's not impossible to do so. First of all, it's crucial to include a control group that gets no drug. Thus, this repeated measures design would include a minimum of 3 levels (No Drug, Drug X, Drug Y). Given that design, you would want to counterbalance the order of the three treatments. With only three levels of the treatment, you would use complete counterbalancing and that would require that you run increments of 6 participants. A reasonable concern about the lingering effects of one drug or the other would also lead you to place some delay between treatments (to allow the effects of a drug to subside).

Given those concerns, I would run 30 rats. I would use each of the 6 orders 5 times. The 6 orders would be: NXY, NYX, XNY, XYN, YNX, and YXN. To ensure that the rats were motivated, I would decrease their food intake for a week. I would inject each rat on Day 1 (with either a placebo, Drug X or Drug Y) and then wait .5 hour for the "drug" to take effect. Then I would place the rat in a maze and record the time for it to complete the maze. Then I would wait a full day and place the rat in a similar (but different) maze after injecting the rat with the next "drug" and waiting .5 hour. The same for Day 3.

7b. Suppose that your data produced the source table seen below. Complete the source table and interpret the results as completely as you can. [5 pts]

ANOVA Table for Drug

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Subject	6	15.238	2.540				
Category for Drug	2	117.714	58.857	36.000	<.0001	72.000	1.000
Category for Drug * Subject	12	19.619	1.635				

There is a significant effect of Drug, $F(2,12) = 36, p < .05$. However, without the means, you would not be able to compute a post hoc test. Nonetheless, you could determine that $HSD = 1.8$. Note that there is a problem with counterbalancing in this study. With $df_{\text{subj}} = 6$, you know that there were 7 rats in the study, which means that the study could not have been appropriately counterbalanced.

7c. Suppose that you redesigned your study as an independent groups design. Suppose, further, that (through some major miracle) you get exactly the same scores as you achieved with your repeated measures study. First of all, tell me how many rats you would have in your independent groups study. Then, produce a source table below to show what your analysis would look like. [5 pts]

Source	df	SS	MS	F
Treatment	2	117.7	58.9	31
Error	18	34.9	1.9	
Total	20	152.6		