1. You've read the Mook article, with its extended discussion of external validity. First, 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]
The answer to this question relies on a fairly straightforward application of information from the Mook article.

2. Define an operational definition and then give an example of a specific operational definition from the Mook article, the Shlenker article excerpt, from the Festinger & Carlsmith study, or the Darley & Latané study. [5 pts]

In general, an operational definition is a clear and concrete statement of the process involved in the definition of a term. As Ray says, "…it redefines the (general) concept in terms of clearly observable operations that anyone can see and repeat. (p. 55)" To measure anxiety, Ray suggests that we might define anxiety from self-report measures, physiological measures, or observation of a specific behavior (hand wringing? ☺). "…we might define aggression as the number of times a child hits a toy… (Ray, p. 55)"

Given the wide range of options, you should be able to find a lot of operational definitions.

3. Why is a nonmanipulated characteristic of a participant not a true independent variable? What is the source of ambiguity in doing research with such nonmanipulated characteristics? [5 pts]

I would argue that such a nonmanipulated characteristic is not a true independent variable because it cannot be manipulated. With a true independent variable, participants can be randomly assigned to levels of the treatment. That cannot be the case for nonmanipulated characteristics of participants (like gender, race, age, etc.). The ambiguity emerges because of the fact that such characteristics tend to be linked to (correlated with) other characteristics. And, thus, there is the option that any observed effect could be due to the other characteristics. (Essentially a third variable problem.) For example, suppose that you were studying age as your "IV." You were using 6-yr olds, 36-yr olds, and 96-yr olds. You should readily see that these participants are likely to vary in many ways in addition to age (e.g., education, motor skills, sensory acuity, etc.).

4. Professor Ty Knott was interested in the relationship between the longevity of divorced women's marriages and the longevity of their divorced parent's marriages. To that end, he collected a sample of 25 divorced women whose parents had also been divorced. The StatView analysis of the data is seen below. Interpret these data as completely as you can. If a woman's parents had been married for 10 years prior to their divorce, how long would you predict that the woman's marriage would last before a divorce? If a woman's parents had never been divorced, what prediction could you make about the length of time before the woman's marriage might end in divorce? [10 pts]

Regression Summary
Yrs. Person Married vs. Yrs. Parents Married

Count	25
Num. Missing	0
R	.852
R Squared	.726
Adjusted R Squared	.714
RMS Residual	1.510

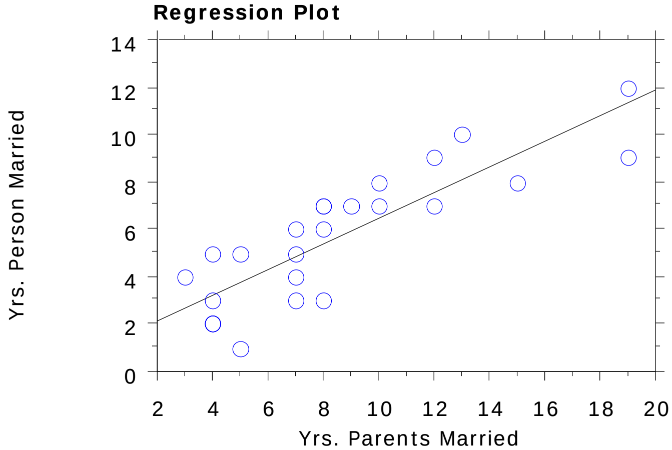
ANOVA Table
Yrs. Person Married vs. Yrs. Parents Married

	DF	Sum of Squares	Mean Square	F-Value	P-Value
Regression	1	139.013	139.013	60.985	<.0001
Residual	23	52.427	2.279		
Total	24	191.440			

Regression Coefficients
Yrs. Person Married vs. Yrs. Parents Married

	Coefficient	Std. Error	Std. Coeff.	t-Value	P-Value
Intercept	1.108	.659	1.108	1.682	.1061
Yrs. Parents Married	.539	.069	.852	7.809	<.0001

Regression Plot



Y = 1.108 + .539 * X; R^2 = .726

With r = .852 and p < .0001, there is a significant positive linear relationship between Years Parents are Married and Years an Individual is Married. (You know it's positive from the positive slope.) With $r^2 = .7$, that means that the two variables share a lot of variability (~71%). Thus, much of the variability in the longevity of a divorced person's marriage is related to the longevity of that person's divorced parent's marriage. You should note, however, the selectivity of the sample (divorced people with divorced parents). [And, of course, the data are totally fabricated. If the parents had been married 10 years prior to divorcing, then you would predict that the person would be married ~6.5 years before divorcing. You wouldn't be able to make any prediction at all if the parents hadn't divorced, because your sample didn't include anyone with that marital background.

5. We discussed the Darley & Latané study in class. Remember, that's the one in which a person is in a booth for a "communication" study, when one of the other "participants" appears to be in physical danger (choking, etc.). The study was designed to test bystander apathy. OK, that should be enough information to allow you to recall the basic study. Well, it's not widely known, but another pair of investigators was also interested in this issue — Dudley & Lamé. They did a study in which the participant thought that there were 5, 10, or 15 bystanders present in a communication study. The participant went into a booth, as did the appropriate number of other "participants." (Each person in a different booth, of course.) The researchers measured the time (in minutes) before the participant would go to the aid of a fellow participant who appeared to be choking uncontrollably. First, interpret the results of this study as completely as you can.
The results of the study were not significant (p > .05), so the three groups did not appear to differ significantly in time to respond. I would argue that the study was not sufficiently powerful, largely because of the absence of a "control" group, as seen below.

Then, answer the following questions:
a. Can you detect any problems with the design of this study? What does it appear to be lacking? How would you redesign the study to correct the problem(s)? Be very explicit.
The major problem is the absence of a "control" group with only the participant and one confederate present (Group = 2). You should certainly add such a group to the study, or replace one of the groups (size = 5, 10, 15) with such a group (size = 2).
b. Dudley and Lamé suggest that they study should be conducted again as a repeated measures design. If you think about the design of the study, you should be able to give them at least one very good reason that the study couldn't be conducted as a repeated measures design. [20 pts]
The nature of the experiment would not lend itself to a repeated measures design. That is, because of the deception, you'd only be able to fool the participant once. (Or, if you could fool the participant more than once, you probably wouldn't want to use that person's data! ☺)

ANOVA Table for Time to Respond

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
No. of Bystanders	2	2.400	1.200	1.761	.1911	3.522	.325
Residual	27	18.400	.681				

Means Table for Time to Respond
Effect: No. of Bystanders

	Count	Mean	Std. Dev.	Std. Err.
5	10	2.800	.632	.200
10	10	3.400	.966	.306
15	10	3.400	.843	.267

6. Dr. June Bugg was interested in examining the extent to which one's attention is captured by a problem. She sets up her experiment on the computer so that the participant is presented with a series of problems, one at a time. Participants are told to solve as many problems as they can within a 10-minute period. At the same time, participants are told that when a small dot appears in the bottom left corner of the computer screen, they should hit the space bar. The experiment lasts about an hour, with 5 different types of problems that vary in difficulty (Very Easy, Easy, Moderate, Difficult, and Very Difficult). Dr. Bugg wants to "warm up" the participants for the Very Difficult problems, so she runs every participant through the Very Easy problems first (for 10 minutes), then through the Easy problems (for the next 10 minutes), then through the Moderate problems (10 minutes), then through the Difficult problems (10 minutes), and finally through the Very Difficult problems (10 minutes). [Obviously, in a 10-minute interval people will solve more easy problems than difficult problems, but that's not the dependent variable.] At 10 random times within each 10-minute period, a dot appears in the bottom left corner of the screen. The dependent variable in this experiment is the number of times that the participant detects the dot at the bottom of the screen. Dr. Bugg reasons that as tasks become more difficult, the participant's attention will be more absorbed by the problem, leading the participant to miss the appearance of the dot. Below is a partially completed source table for this experiment. Interpret the results as completely as you can. [20 pts]

ANOVA Table for Task Difficulty

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Subject	11	20.733	1.885				
Category for Task Difficulty	4	145.500	36.375	88.425	<.0001	353.702	1.000
Category for Task Difficulty * Subject	44	18.100	.411				

Means Table for Task Difficulty
Effect: Category for Task Difficulty

	Count	Mean	Std. Dev.	Std. Err.
Very Easy	12	9.333	.778	.225
Easy	12	9.083	.669	.193
Moderate	12	8.250	.622	.179
Difficult	12	7.500	.522	.151
Very Difficult	12	5.000	1.348	.389

The very first point that you should note is that in this repeated measures design, conditions are not counterbalanced. That's a major confound, which makes the results impossible to interpret cleanly. (Are the effects due to difficulty or simply due to fatigue?)

To complete the analysis, you would need to compute HSD, which in this case would be .75 (q = 4.03, with 5 treatments and 44 df). Thus, were it not for the confound, you could conclude that people missed the dot significantly more often on the VD problems compared to all other problems. They missed the dot significantly more often on the D problems compared to all other problems. They missed the dot significantly more often on the M problems, compared to E and VE problems. However, they missed the dot equally often on E and VE problems.