SYMSYS 130: Research Methods in the Cognitive and Information Sciences (Spring 2013)

Take Home Final
June 20, 2013 – Instructor' s Responses

Please respond to the following questions with short essays (300-500 words, not more). Answers will be scored out of 25 points total, based on the following criteria (5 pts each):

- informativeness (interesting, nonobvious),
- correctness (sound, accurate),
- thoroughness (convincing, rigorous),
- coherence (consistent, well constructed), and
- conciseness (clear, succinct).

1. In a Massive Open Online Course (a "MOOC"), students have access to an official course message board/forum, also used by the TAs and course instructor. Without doing a randomized controlled experiment, how would you test the hypothesis that participating in the forum improves a student's performance in the course? How does your design control for potential confounding variables?

I would begin by testing whether participation in the forum correlated positively and significantly with a student's performance. Performance could be measured by GPA as the dependent variable (DV), but I would exclude any component based on forum participation, e.g. if students received participation points for forum postings, this would need to be separated from the DV. A possible DV is performance on a final exam or other measure into which forum participation did not enter. For rigor, we could talk to the graders to see if their scoring might have been influenced positively by recognizing that a student had participated in the course forum. Assuming a pure DV measuring performance, we would define an independent variable measuring forum participation. The most comprehensive measure would include both postings and readings of the forum, since some students might read the forum without posting. Separate measures for number of comments read and number of comments posted could be combined into a third measure for total participation, by adding comments read to comments posted. These would be the independent variables or IVs.

We could then see whether the correlation was positive between each IV and the DV. Positive correlations that were statistically significant would suggest an effect for forum participation on an independent measure of student performance, but it would not rule out possible confounds. A student's interest in, talent for, and/or knowledge of the material in the course could motivate them to use the forum more than other students and alsoo cause them to perform better. To control for these confounds, since we cannot do a controlled experiment, there are a few possibilities and I would ideally try both of them:

(a) Gather data on performance for each student who took the MOOC and also took a similar course that did not have course forums, on the assumption that a student's interest, talent, and knowledge would be consistent across the different courses. This would restrict the comparison to the subset of students who took both courses. If students in the upper 50% on forum participation in the target showed a larger grade or exam improvement relative to the similar course than did those in the lower 50% on the IVs, this would be evidence that forum

participation and not one of the confounding variables boosted the performance of the students who participated in it heavily.

(b) Poll students in the course to assess their interest, talent, and knowledge levels for both the target course and the similar course mentioned in (a) above, using self-rating. This would check the extent to which confounding variables were indeed similar across the two courses for individual students. It would also allow us to do a multiple regression model in which the poll questions were treated as additional IVs along with those for forum participation, to see whether forum participation showed an effect over and above any correlation it might have had with the potential confounds.

2. Suppose there are two paired variables x and y which are both binary (they assume values 0 or 1), and that we have a list of the paired values of x and y for an index from 1 to N. (a) How would you calculate Pr(x=1|y=0) for these data? (b) Express the Pearson correlation coefficient between x and y as a function of the joint probabilities for x and y.

(a) $Pr(x=1|y=0) = Pr(x=1 \ \& \ y=0) / Pr(y=0)$, so to calculate our target probability, we add up all the pairs for which x=1 and y=0, and divide by those for which y=0. $Pr(x=1|y=0) = [\#\{i: x_i=1 \ \& \ y_i=0\}/N]/[\#\{i: y_i=0\}/N] = \#\{i: x_i=1 \ \& \ y_i=0\} / \#\{i: y_i=0\}$.

(b) From Trochim, the formula for the Pearson correlation coefficient is

$$r = [N\sum xy - (\sum x)(\sum y)] / \text{sqrt}\{[N\sum x^2 - (\sum x)^2][N\sum y^2 - (\sum y)^2]\}.$$

Multiplying the right hand side by $(1/N^2)/(1/N^2)$ yields

$$r = \{[\sum xy - (\sum x)(\sum y)]/N^2\} / \text{sqrt}\{[N\sum x^2 - (\sum x)^2][N\sum y^2 - (\sum y)^2]\}/N^2$$
$$= \{\sum xy/N - [(\sum x)/N][(\sum y)/N]\} / \text{sqrt}\{[N\sum x^2 - (\sum x)^2][N\sum y^2 - (\sum y)^2]\})/N^4\}$$
$$= [E(xy) - E(x)E(y)] / \text{sqrt}(\{E(x^2)-[E(x)]^2\}\{E(y^2)-[E(y)]^2\}) \quad \text{(FORMULA 1)}$$

Let *x* denote x=1, and *y* denote y=1, so that ~*x* denotes x=0 and ~*y* denotes y=0. Let $p(x,y) = Pr(x=1,y=1)$. Then $p(x) = Pr(x=1) = 1-Pr(x=0) = 1-p(\sim x)$, and $p(y) = Pr(y=1) = 1-Pr(y=0) = 1-p(\sim y)$. We can solve for the terms in FORMULA 1 as follows:

$$E(xy) = p(x,y)(1*1) + p(x,\sim y)(1*0) + p(\sim x,y)(0*1) + p(\sim x,\sim y)(0*0)$$
$$= p(x,y).$$

$E(x) = p(x)(1) + p(\sim x)(0) = p(x)$, and analogously $E(y) = p(y)$.

Furthermore,

$$\text{sqrt}(\{E(x^2)-[E(x)]^2\}\{E(y^2)-[E(y)]^2\}) = \text{sqrt}(\{E(x^2)-[E(x)]^2\}\{E(y^2)-[E(y)]^2\})$$
$$= \text{sqrt}(\{[p(x)*1^2+p(\sim x)0^2] - [p(x)*1 + p(\sim x)*0]^2\}\{[p(y)*1^2+p(\sim y)0^2] - [p(y)*1 + p(\sim y)*0]^2\})$$
$$= \text{sqrt}[p(x)p(\sim x)p(y)p(\sim y)], \text{ which equals the product of the standard deviations for x and y.}$$

Therefore,

$$r = [p(x,y) - p(x)p(y)] / \text{sqrt}[p(x)p(\sim x)p(y)p(\sim y)].$$

[NOTE: This is equivalent to the phi coefficient and to the Matthews correlation coefficient for paired binary variables.]

3. Describe how you would design a phone app to help improve people's ability to overcome functional fixedness in problem solving. Describe how your knowledge of human psychology informs the design.

There are many possible designs. One idea would be to create a multiplayer game app called “Find the Uses” in which users, prompted by a name or picture of an object, would type in a use for that object. If the typed-in name matches an approved use in the app's central

database, the user would get 1 point. If there is no match, then the user's suggested use would be put to a random subset of other users for a vote. Voters would get to see the list of already approved uses for the object, and would vote whether the suggestion was (a) a novel use, (b) a paraphrase of an existing use, or (c) not a legitimate use of the object. The user would get 1 point for an automated match. Otherwise the user would get 5 pts if a majority of users vote it a novel use, 1/2 pt if a majority votes it a paraphrase, and 0 pts if there is no majority. Top point getters will get recognized in the app. This design takes advantage of the fact that people like to play games for virtual points and to get recognized, and it aligns their incentives so that learning how to think of novel uses of an object will cause them to do better in the game. Learning happens through practice, and this game should instill the habit of thinking of novel uses of an object, hence overcoming the functional fixedness bias in problem solving.

4. Provide a practical principle for determining whether a particular application of neuroscience to marketing is ethical. Make an argument that this principle would be likely to be adopted, how it would be effective, and why it would result in improved human welfare?

An example of such a principle might be the following:

“Any use of neuroimaging technology in marketing research that affects the design of the product or the content of advertising for the product should be disclosed using standard language that is approved by the National Institute of Mental Health (NIMH).”

This is similar to other warnings such as the Surgeon General's warning on cigarette packages and ads and mandated warnings about the side effects of FDA-approved drugs. Therefore, and also because there is widespread concern about the mind-controlling potential of neuromarketing, it is plausible that Congress would adopt such a requirement, which would lead to its widespread adoption by companies that want to avoid running afoul of the law. Such mandatory disclosure would be effective in the sense that it would (a) make people more aware of the presence and extent of neuromarketing and (b) encourage them to take this into account when making purchases of affected products. I believe human welfare would be improved by this because more people would think independently about whether they really want a product, taking into account the fact that they have been neuromarketed. This should lead to deeper thought about how much value the product would add to one's life, avoiding some purchases whose value, on reflection, is not justified by their expense.

5. Choose a paragraph from any of the assigned reading this quarter, and rewrite it so that it communicates its intended points more effectively.

Consider the following paragraph from Coombs, Dawes, and Tversky, which was the subject of a question in one of the homework assignments in this course:

“A more difficult problem arises with respect to statements involving numerical values for which no explicit measurement model exists. The measurement of intelligence is a case in point. To justify the use of averages, some psychologists have argued that intelligence is measured on an interval scale. Others claimed that the IQ scale is essentially ordinal and that hence no averaging can be justified. A closer examination of the problem reveals that no measurement theory for intelligence is available. Consequently, no representation theorem can be established and no meaning can be given to the uniqueness problem. This does not

imply that IQ scores are useless. On the contrary, they may provide an extremely useful and highly informative index. In the absence of a well-defined representation relation, however, the uniqueness problem is not well defined."

Here is a rewrite, with the new text highlighted in italics:

"A more difficult problem arises with respect to statements involving numerical values for which no explicit measurement model exists. The measurement of intelligence is a case in point. To justify the use of averages, some psychologists have argued that intelligence is measured on an interval scale. Others claimed that the IQ scale is essentially ordinal and that hence no averaging can be justified. A closer examination of the problem reveals that *there is no agreed definition or objective way to distinguish the intelligence level or type of one person vis-a-vis another. Hence,* no measurement theory for intelligence is available. Consequently, no representation theorem can be established and no meaning can be given to the uniqueness problem. This does not imply that IQ scores are useless. On the contrary, they may provide an extremely useful and highly informative index. In the absence of a well-defined representation relation, however, the uniqueness problem is not well defined."

The above rewrite clarifies the reason for a key claim (about IQ and intelligence) in relation to concepts defined previously.

6. Choose a philosophical claim that you find to be questionable, and design an experiment that would help to test it.

Toward the end of the Chomsky-Foucault debate, Chomsky says the following:

"...if we have the choice between trusting in centralised power to make the right decision in that matter, or trusting in free associations of libertarian communities to make that decision, I would rather trust the latter. And the reason is that I think that they can serve to maximise decent human instincts, whereas a system of centralised power will tend in a general way to maximise one of the worst of human instincts, namely the instinct of rapaciousness, of destructiveness, of accumulating power to oneself and destroying others. It's a kind of instinct which does arise and functions in certain historical circumstances, and I think we want to create the kind of society where it is likely to be repressed and replaced by other and more healthy instincts."

This is an argument in favor decentralized decisions by smaller libertarian communities, based on the claim that they are better than centralized decision making at maximizing "decent human instincts." An experiment that would test this would be to randomly assign two groups of 1000 people for some period (say, one year), to either a centralized or a decentralized environment. In the centralized condition, participants would be instructed and facilitated to elect a single leader and perhaps a council who would govern this temporary society, while in the decentralized one, participants would be instructed and facilitated to make associations with any groups of other participants they choose to, and these free associations would not be empowered to force anything on any group, but could suggest ideas that would then need the consent of any affected community to be put into practice. After the year, we would look at things like whether those in power in the centralized society acted more rapactiously (i.e. used their position to accumulate for themselves rather than serving the general good), whether they hurt others, etc., by a survey comparison with

participants in the decentralized condition.