The outcome is the trustworthiness rating of the salesman. The explanatory variable is the salesman’s race. Each subject is randomly assigned to one of the two treatments.

**Interpretability**
There is no “control” race and no subject blinding or placebo treatment is possible. Blinding for the researcher is not important because the treatment application is just turning on the video and the measurement is fully objective (reading the subjects response). The randomization should reduce confounding factors, such as age, among subjects in the two treatment groups. But, there may be major confounders affecting interpretation of the results: the treatments probably differ in more ways than just race, and some of these will probably affect the outcome. E.g., the two salesmen may differ in height, weight, rapidity of speech, clothing, etc. A better design would at least have several actors of each race randomly chosen to be presented to the subject.

**Generalizability**
Certainly restricting subjects to people who ride the bus and are willing to participate for $5 limits the generalizability of the study. Such people, e.g. will tend to have a lower mean income than the general population.

**Power**
Blocking by subject race will probably result in a smaller error for each subject block, thus increasing power (not to mention the fact that there may well be interaction between subjects race and the main treatment). The subject variability could be reduced further, e.g. restricting to one gender would improve power, but reduce generalizability. The environment could be held constant, e.g. by assuring that the video playback room is kept at a comfortable temperature and free of extraneous noise and visuals. The treatment application should be kept as constant as possible, e.g. making sure that subjects are the same distance from the screen. The measurement seems to provide little room for reducing variability, but perhaps averaging several different trustworthiness scales could be done, and that would reduce measurement variability. A within subjects design, with each subject seeing both salesmen, would improve power (but the presentation order should be randomized). A possible covariate to measure (and add to the model) includes some general measure of how the subject feels about salesmen. Depending on the details of the investigators’ reasons for doing the study, assuring that the actors are easily and clearly identified with a racial group by simple observation might improve the “strength” of the treatments. And always we can improve power by using more subjects.