Within-subject designs in social priming - my attempts

JACOB JOLIJ

TL;DR summary: it's perfectly possible to do a within-subject design for 'social' priming.

This is going to be an attempt at a more serious post, about some actual research I have done. Moreover, I really need to get back into writing mode after summer leave. Just starting cold turkey on the >7 manuscripts still waiting for me did not work out that well, but maybe a nice little blog will do the trick!

This weekend, I was engaged in a Twitter exchange with Rolf Zwaan, Sam Schwarzkopf, and Brett Buttliere about social priming (what else? Ah, psi maybe!) A quick recap: social (or better: behavioural) priming refers to the modification of behaviour by environmental stimuli. For example, washing your hands (and thus 'cleaning your conscience') reduces severity of moral judgments. Reading words that have to do with elderly people ('bingo', 'Florida') makes you walk slower. Or, feeling happy makes you more likely to see happy faces.

The general idea behind such effects is that external stimuli trigger a cascade of semantic associations, resulting in a change in behaviour. 'Florida' triggers the concept 'old', the concept 'old' triggers the concept 'slow', and if you think about 'slow' this automatically makes you walk slower. Indeed, semantics are closely tied to information processing in the brain – a beautiful study from the lab from Jack Gallant shows that attention during viewing of natural scenes guides activation of semantically related concepts. However, whether the influence of external stimuli and semantic concepts is indeed so strong as some researchers want us to believe is questionable. Sam Schwarzkopf argued in a recent blog post that if we are so strongly guided by external stimuli, our behaviour would be extremely unstable. Given the recent string of failures to replicate high-profile social priming studies, many researchers have become very suspicious of the entire concept of 'social priming'.

What does not exactly help is that the average social priming study is severely underpowered. People like Daniel Lakens and Uli Schimmack have done far better in explaining what that means than I can, but basically it boils down to this: if you're interested in running a social priming study (example courtesy of Rolf Zwaan), you pick a nice proverb (e.g., 'sweet makes you sweeter'), and come up with an independent variable (your priming manipulation, e.g. half of your participants drink sugary water; the other half lemon juice) and a dependent variable (e.g., the amount of money a participant would give to charity after drinking the priming beverage). I've got no idea whether someone did something like this… oh wait… of course someone did.

Anyway, this is called a 'between subject' design. You test two groups of people on the same measure (amount of money donated to charity), but the groups are exposed to different primes. To detect a difference between your two groups, you need to test an adequate number of participants (or, your

sample needs to have sufficient power). How many is adequate? Well, that depends on how large the effect size is. The effect size is the mean difference divided by the pooled standard deviation of your group, and the smaller your effect size, the more participants you need to test in order to draw reliable conclusions. The problem with many social priming-like studies is that participants are only asked to produce the target behaviour once (they come into the lab, drink their beverage, fill out a few questionnaires, and that's it). This means that the measurements are inherently noisy. Maybe one of the participants in the sweet group was in a foul mood, or happened to be Oscar the Grouch. Maybe one of the participants in the sour group was Mother Theresa. Probably three participants fell asleep, and at least one will not have read the instructions at all.

To cut a long story short, if you don't test enough participants, you run a large risk of missing a true effect (a false negative), but also you risk finding a significant difference between your groups whilst there is no true effect present (a false positive). Unfortunately, many social priming studies have used far too few participants to draw valid conclusions. This latter thing is significant (no pun intended). Given that journal editors until recently primarily were interested in 'significant results' (i.e., studies that report a significant difference between two groups), getting a significant result meant 'bingo'. A non-significant result… well, too bad. Maybe the sweet drink wasn't sweet enough to win over Oscar the Grouch! Add a sugar cube to the mix, and test your next batch of subjects. If you test batches of around 30 participants (ie. 15 per group, which was not abnormal in the literature), you can run such an experiment in half a day. Sooner or later (at least within two weeks if you test full time), there will be one study that gives you your sought-after p < .05. Boom, paper in!

In cognitive psychology and neuroscience we tend to be a bit jealous of such 'easy' work. Our experiments are harder to pull off. Right before summer, one of my grad students finished her TMS experiment, for which she tested 12 participants. For 18 hours. Per participant. In the basement of a horrible 1960s building with poor airco whilst the weather outside was beautiful. Yes, a position in my lab comes with free vitamin D capsules because of occupational health & safety reasons.

Moreover, the designs that we typically employ are within-subject designs. We subject our participants to different conditions and compare performance between conditions. Each participant is his/her own control. In particular for physiological measurements such as EEG this makes sense: the morphology, latency and topography of brain evoked potentials vary wildly from person to person, but are remarkably stable within a person. This means that I can eliminate a lot of noise in my sample by using a within-subjects design. As a matter of fact, the within-subjects design is pretty much the default in most EEG (and fMRI, and TMS, and NIRS, etc.) work. Of course we have to deal with order effects, learning effects, etc., but careful counterbalancing can counteract such effects to some extent.

Coming from this tradition, when I started running my own 'social priming' experiments, I naturally opted for a within-subjects design. My interest in social priming comes from my work on unconscious visual processing – very briefly, my idea about unconscious vision is that we only use it for fight-or-flight responses, but that we otherwise rely on conscious vision. The reason for this is that conscious vision is more accurate, because of the underlying cortical circuitry. Given that (according to the broad social priming hypothesis) our behaviour is largely guided by the environment, it is important to base our behaviour on what we consciously perceive (otherwise we'd be acting very odd all the time). This led me to hypothesize that social priming only works if the primes are perceived consciously.

I tested this idea using a typical masked priming experiment: I presented a prime (in this case, eyes versus flowers, after this paper), and measured the participant's response in a social interaction task after being exposed to the prime, in total 120 trials (2 primes (eyes/flowers) x 2 conditions (masked/not masked) x 30 trials per prime per condition). The 'social interaction' was quite simple: the participant got to briefly see a target stimulus (happy versus sad face), and had to guess the identity of the face, and bet money on whether the answer was correct. Critically, we told the participant (s)he was not betting her/his own money, but that of the participant in the lab next door. Based on the literature, we

expected participants to be more conservative on 'eye'-primed trials, because the eyes would remind them to behave more prosocially and not waste someone else's money.

Needless to say, this horrible design led to nothing. Major problem: it is very doubtful whether my DV truly captured prosocial behaviour. After this attempt, we tried again in a closer replication of earlier eye-priming studies using a between-subjects design and a dictator task, but after wasting >300 participants we came to the conclusion many others had drawn before: eye priming does not work.

But this doesn't mean within-subjects designs cannot work for priming studies. There's no reason why you could not use a within-subjects design to test, for example, whether having a full bladder makes you act less impulsively. As a matter of fact, I've proposed such a study in a blog post from last year.

Another example: I am not sure if we could call it 'social priming', but a study we did a while ago used a within-subject design to test the hypothesis whether happy music makes you better at detecting happy faces and vice-versa. Actually, this study fits the bill of a typical 'social priming' study – activation of a high level concept (happy music) has an effect on behaviour (detecting real and imaginary faces) via a very speculative route. It's a 'sexy' topic and a finding anyone can relate co. It may not surprise you we got a lot of media attention for this one…

Because of the within-subjects design we got very robust effects. More importantly, though, we have replicated this experiment two times now, and I am aware of others replicating this result. As a matter of fact, we were hardly the first to show these effects… music-induced mood effects on face perception had been reported as early as the 1990s (and we nicely cite those papers). The reason I am quite confident in the effect of mood on perception is that in our latest replication, we also measured EEG, and indeed find an effect of mood congruence on visual evoked potentials. Now, I am not saying that if you cannot find a neural correlate of an effect, it does not exist, but if you do find a reliable one, it's pretty convincing that the effect *does* exist.

What would be very interesting for the social priming field is to come up with designs that show robust effects in a within-subjects setting, and ideally, effects that show up on physiological measures. And to be frank, it's not that difficult. Let's suppose that elderly priming is true. If concepts related to old people makes you indeed behave like grandpa, we should not just see this in walking speed, but also in cognitive speed. Enter the EEG-amplifier! Evoked potentials can be used nicely assess speed of cognitive processing – in a stimulus recognition task, for example, latency of the P3 correlates with reaction time. If 'old' makes you slower, we'd expect longer P3 latencies for trials preceded by 'old' or a related word, than for trials preceded by 'young'. Fairly easy experiment to set up, can be run in a week.

Or even better – if, as the broad social priming hypothesis postulates, social priming works by means of semantic association, we should be able to find evidence of semantic relations between concepts. Again something that is testable, for example in a simple associative priming task in which you measure N400 amplitudes (an index for semantic relatedness). As a matter of fact, we have already run such experiments, in the context of Erik Schoppen's PhD project, with some success – we were able to discriminate Apple from Android enthusiasts using a very simple associative priming test, for example.

All in all, my position in the entire social priming debate has not changed that much. I do believe that environmental stimuli can influence behaviour to quite some extent, but I am very skeptical of many of the effects reported in the literature, not in the last place because of the very speculative high-level semantic association mechanisms that are supposed to be involved. In order to lean more credibility to the claims of 'social priming', the (often implicit) hypotheses about the involved mechanisms have to be tested. I think we (cognitive/social neuroscientists) are in an excellent position to help flesh out paradigms and designs that are more informative than the typical between-subject designs in this field. At least I think working together our colleagues in social psychology this way is more fruitful than trying to 'educate' social priming researchers about how 'wrong' they have been, and doing direct replications

(however useful) of seminal studies and bask in Schadenfreude when the yet another replication attempt fails, or meta-analysis shows how flimsy an effect is. We know that stuff already. No need to piss each other off IMO (I am referring to a rather escalated ISCON discussion here of last week).

Let's do some cool stuff and learn something new about how the mind works. Together. Offer made last year still stands.