Summary: A single experimental simulation of intelligence gathering from interviews with members of a mock network is proposed, as a test of the disclosure-outcomes management (DOM) model, and to inform practitioners on appropriate interview strategies for intelligence gathering from network members. In phase 1, participants interact in groups of 6 to plan their intended information revelation. In phase 2, individuals go through an online ‘interview’ comprising a sequence of interviews and selections from a drop-down menu. The extent to which their information sharing is accounted for by mixed effects random models of individual and group presents their approach to testing whether individuals make their own disclosure decisions or are guided by the group plan…. I think (-:

Recommendation: This is a fascinating area to explore, and an ingenious approach. The study has been designed with precision and the analysis plan looks sound. It is a very complex design indeed, and I am not 100% sure I fully understand it. My main concern is about the practical relevance of the research design to the real task of intelligence gathering. I acknowledge it is exceptionally difficult to study (see, for example, our own struggles: Dando & Ormerod, 2019 – Journal of Experimental Psychology: General). But I do worry that the design is so convoluted and potentially confounded, and that the task set is so far from anything that would realistically be encountered in an intelligence-gathering setting, that the efforts in data collection may end up providing more noise than signal. I raise a set of issues below that I think need addressing. I do hope a version of this study can be accepted, but I also think the issues below require attention or a robust dismissal before this pre-registration can be accepted.

Hypotheses:

My understanding of the predictions of DOM are summarised in this table:

|  |  |  |
| --- | --- | --- |
|  | Low benefit | High benefit |
| Low risk | “low stakes” – no reveal | “Unguarded” - reveal |
| High risk | “Guarded” – no reveal | “High stakes” – maybe reveal |

Re. Q1 – The DOM predictions seem fairly facile to me. What theory would predict otherwise? I think this needs to be addressed explicitly for the value of the study to be justifiable.

P3 “When interviewing someone about their network—and the focus is solely on the network—to what extent does network membership predict the *type of information* the interviewee will *choose* to reveal. Are two or more different people from the same network likely to disclose similar kinds of information? Or does disclosure in this context better resemble individuals independently managing the potential outcomes of revealing information?”

Aren’t all the ‘networks’ essentially identical, so how could this prediction be tested? Surely it would be a much more powerful study if the underlying structure of networks was manipulated (e.g., a hierarchy vs. a heterarchy) so that network influences could be detected with precision?

P5” Such disclosure is not immediately beneficial to navigating the self-interest dilemma but carries potential costs. “

I thought the point of low-stakes information was that it doesn’t carry potential costs. The DOM model does not therefore predict low levels of disclosure – one would need a theory of verbal manoeuvring such as Interpersonal control theory to derive this prediction.

P6 – “result a” vs “result b” – The comparison between results a and b seems to boil down to this: if this research tells us that everyone in a network will be influenced by the same cost/benefit judgements, then they will all say the same thing in interview, so you can interview them all in the same way. If not, they won’t, so you have to treat each individual differently.

I don’t understand how this informs an interview questioning approach. The detective identifies from another source (a tip off) what would constitute a high stakes piece of information. She/he then interviews one member network who reveals that high stakes piece of information. So the detective thinks ‘right, I’ll repeat the same interview approach because that revealed the high stakes information’. But a) they already knew the high stakes information, and b) they had it confirmed in the first interview, c) knowing what is high or low stakes isn’t the thing that is helping the detective: what is helping is the interview method they chose for the first interview. Perhaps most importantly, d) the idea that one can draw any kind of generality from this research to interviewing members of a real network, where there are different roles, hierarchical positions, motivations, fears, intellectual competencies, personalities, etc. among members, seems wrong. To advise any detective to ignore individual characteristics in planning an interview approach would seem to be very poor advice indeed.

P7.” We aim to recruit already acquainted participants—for example, friends or co-workers—to serve as a network: typically, network members are not complete strangers. This design choice allows us to commence test sessions without needing to induce familiarity between participants”

This decision imposes a limitation on the generalisability of the results. a) Doesn’t this bias the results towards intra-network agreement (i.e., a confirmation bias on the predicted results)? B) It doesn’t necessarily reflect the reality of networks (e.g., Al Quida is a brand not a network of friends), and c) it introduces huge sources of variability (notably experience and trust) that do not form part of DOM theory.

P7. They propose 20 networks of 6 individuals in each. Is this 20 x 6 = 120 or just 20 data points? Note on P3 they state “We will explore how perceived disclosure outcomes, namely, the projected costs and benefits, affect what networks choose to reveal. The study will focus on disclosure pertaining to the network as a whole, not about the individual being interviewed, per se” So is it networks or individuals? I think its individuals not networks, so this quote from p.3 really needs to be adjusted.

P9 - “They will execute their decisions by selecting what to disclose from a list of possible information items”

Doesn’t this create a massive demand characteristic to disclose (as opposed, say, to a ‘no comment’ interview, a much more realistic outcome of an intelligence interview)?

P9. This took a lot of understanding, and I am not sure I do understand it. At first I thought “Doesn’t the ‘probability’ information effectively tell participants what they can and cannot reveal? So, they will simply maximise release of low risk and minimise high risk”. But having got the whole way through, it seems that the probability information is only marginally predictive of the actual gains and losses. That said, it seems like the participant only discovers that once they have had the second video (i.e., the first trial). So, in all likelihood, participants are going to treat these probabilities as if they are veridical rather than pseudo-random in their informativeness.

Providing the % safety information during the interview seems a bit odd – presumably the group will discuss what they think is safe before, and then this % info comes along that may override or confirm the group discussion… how does that work? Is the idea that the % info should conflict with the group decision? To be realistic, the networks ought to decide what is high and low stakes, and the individuals when interviewed should be left to make their own judgements on the basis of what was planned by the network, and not be told what is high and low stakes by the interviewer. The proposed approach seems to take the study farther away from anything that has practical validity.

P9 – “We will tell them that they can double their endowment in the best case, and in the worst case, they can lose the entire endowment.”

There seem to be competing dynamics here: on the one hand, the study is looking at group cohesiveness and information disclosure, bit on the other hand it seems to be a behavioural economics study looking at how well individuals can maximise profit and minimise losses. I worry that this financial incentive creates an entirely different set of task goals to the ones the research team think they are studying. All the stuff about keeping/losing licences, stopping/continuing investigations and disclosing some information to the interviewer to appear cooperative is made irrelevant, or at very least is confounded, by this financial incentive.

P10. Is 20 mins planning enough? Where does this figure come from?

P11. “In truth, each participant will be told that the investigators have called them in for an interview.”

If the participants are members of an existing friendship group interviewed over Zoom, how will the researchers avoid collusion and information sharing?

P.11 Should level of affiliation questionnaire come after the interview not before, so as not to bias the data?

p.11 Surely presenting % safe and %dangerous estimates is two views of the same measure If its 10% safe, doesn’t that necessarily mean its 90% dangerous? This scale is not the same thing as indicating costs and benefits: the cost (e.g., in terms of risk of incrimination) can be entirely independent of the benefits (e.g., showing cooperation with the interviewer), whereas safe vs. dangerous measures are clearly not independent.

p.11. I don’t really understand the bit about “disclosing a piece of information will bring about a random outcome” – I am guessing that it means when a participant gives something away, they are told how much cost/benefit in terms of cash that generated. But its not very clear. Having been through the Qualtrics link, I see now that the link between the %s selected and the fiscal outcome is what is randomised. This is really odd for the participant. For example, on trial 1, I selected 15%/15% safe/dangerous figures and ended up with a net loss of 2.5 SEK. On trial 2 I went for 50%/50% and ended with a net gain of 5SEK. However, I don’t think I can have understood this correctly, since on p. 14 the researchers state that the gains and losses with each combination are non-random. Whatever the meaning of this, I worry that there will be unintended carryover between trials on the basis of this trial-by-trial feedback that will interfere with any measure of the influence of network-level planning.

p.13 – what is a ‘memory check’ in this context? It needs to be specified clearly, since it might confound the underlying incentive/task structure.

p.13 – what is described as an ‘interview’ is not in any meaningful sense an intelligence interview. Participants are simply asked to select from a list their responses to three video segments. This lack of realism is highly limiting for the generalisability of the results. I appreciate the researchers have considered this and have made a case for this being a starting point, but I do wonder whether it really tells us anything meaningful about information disclosure during intelligence interviews rather than simply information disclosure, per se.

NB. Shouldn’t the relationship with Prisoners dilemma/game theory be made explicit?