Response Pollet - 016.Vidi.175.343

I am thankful and honoured that all 3 reviewers (R's) praised my academic achievements. Moreover, the outlined research was positively reviewed (A+/A+/A-B). Having said this, the reviewers, in particular R3, also raised some relevant issues and I am grateful for the opportunity to deal with these points in more detail here. As at least two of the R's themselves indicate, none of their issues raised touch on the core of the outlined research programme. Given that more substantial points were raised by R3, and to a lesser extent R2, this response is structured accordingly. Needless to say, I appreciate **all** suggestions, and thank the reviewers for their constructive comments.

R1.

R1 is unequivocally positive across all dimensions, but indicates a potential improvement under 'Final assessment'(c). In particular, R1 suggests that apart from investigating structural factors, the project can be expanded to examining neighbourhood social factors (e.g., social cohesion/trust). This is an excellent suggestion and where available I will include these factors.

R2.

R2 lists two suggestions under and knowledge utilization(b): **(1)** extending to the theoretical underpinnings for the mechanisms at multiple levels and **(2)** considering different forms of prosociality in different contexts, operating via different mechanisms (generalizability).

1. I fully agree with R2 that one should consider the mechanisms at play, most notably the role of salience, he/she suggests (also covered in ref. 20 in the proposal). That being said, I did already consider certain mechanisms, for example, a cultural learning mechanism, lies at the heart of Study 5, where the role of acculturation as a mechanism is explicitly tested. Moreover, the proposed set of studies present a necessary first step to uncovering the (psychological) mechanisms proposed. Having said this, to fully examine all possible mechanisms at play, one ideally needs large-scale, longitudinal data over an entire life course. Within the scope of a VIDI, it would be very hard to collect such data. Yet, relevant secondary data on potential mechanisms are available from various sources (e.g., the National Child Development Study / Millennium Cohort Study), but these do not have the Dictator (DG) measure. Thus, after establishing childhood and neighbourhood SEP effects with the DG, I argue that the project can be expanded in such a direction. However, before closely examining the full range of potential mechanisms via which childhood SEP and neighbourhood SEP effects operate, and theorizing about those, we need to first establish that neighbourhood and childhood SEP do **actually** influence prosocial behaviour: It is this particular knowledge gap that my proposal addresses.

2. This comment is interrelated to the above point, with R arguing for a finer-grained situational approach. In particular R suggests that it is necessary to reflect on the degree to which results based on the DG can be generalized. For example, the

project can be easily expanded to investigate (for example via use of the lost letter method, ref.114 in proposal).

In conclusion, the project can be logically expanded in the directions suggested by R.

R3.

R3 raises three methodological points: **(1)** statistical power, **(2)** selectivity bias, **(3)** the use of the DG in particular with online samples. Finally, the reviewer suggests my interdisciplinary background as a potential risk for the project.

1. In the Bayesian framework I adopted, (frequentist) statistical power is rarely presented, as it is based on *p*-values, which is [problematic](). Traditional power analysis is further complicated by the non-normal distributions and the multilevel designs. Therefore, I explicitly chose not to present traditional power analyses. However, the sample sizes have been argued for throughout the paper, contrary to R's claim. For example, the number of countries and cities are based on a simulation study cited in the original proposal ([ref.117]()). Throughout, the samples are based on previously published studies, a simulation study (ref.117), and where doubt exists, I proposed piloting the procedure (via [simR]() crude estimations can be obtained to plan the exact number). Finally, I would like to underline that the proposed samples are well in line with previous work (refs.). Moreover, apart from power, other factors (such as cost) should be considered. Obviously if there are opportunities to collect more data (e.g., via my faculty's equipment), I will seize these opportunities but as it stands the proposed minimum sample sizes are, contrary to the referee's claim, well-argued for and cost-efficient.

2. R raises the issue of selectivity bias as a consequence of mobility, and the ability to make causal claims as a consequence. In particular, he/she points to the ***mechanism*** of mobility between neighbourhoods to account for neighbourhood effects. Firstly, I would note that the longitudinal nature of Study 5, can address in part the proposed selection effects. Perhaps surprisingly some novel methods might also address the ***plausibility of causality***. These include ['ecological inference'](), and [Bayesian SEM]()/ [networks]() / [directed acyclic graphs](). Therefore, I am optimistic that some scenarios can be shown to be more plausible, even if selectivity bias occurs. For example, one could model how large mobility effects would have to be to account for the observed effects. Finally, and most importantly, R is tacitly ***assuming*** that there is a neighbourhood effect. Note that currently the research gap is not what the mechanism is via which neighbourhood effects occur but rather whether there is ***any neighbourhood effect at all***. As a whole, this set of studies will provide invaluable insights on the degree to which different levels can account for variation in prosocial behaviour (as acknowledged by R1-R2). I fully share the concern about potential mechanisms, and as discussed under R2, the project can be expanded in that direction to further examine those but before is establishing the degree to which different levels can account for the observed patterns.

3. The reviewer questions the use of the DG measure and particularly the use of this

measure with online samples. Firstly, it is important to highlight that this comment regarding the online samples, **only** applies to studies 1-3 rather than the entir proposal. My primary reason for using such samples is that these are commonly used in prosocial behaviour research (e.g., refs. 92,100-101) thus ensuring comparability between prior work and my own results. Thus, in contrast to what the referee suggests, such samples have been widely used and form an established method (also see refs. 37-38 & 91). Moreover, the costs of these samples are substantially lower compared to non-crowdsourced alternatives, such as [Flycatcher](). I am able to apply to my university's equipment fund to pay for the use of Flycatcher, as I do feel the crowdsourced samples could be complemented in this way. However, I consider the use of crowdsourced populations to be actively beneficial in this case, as it ensures comparability with existing datasets, and the ability to make such comparisons is an explicit goal of the project. Finally, it should be noted that this criticism about **only** applies to a subset of the studies and can be remedied. Crowdsourcing comparability with previous work and because of its lower costs, but complementary samples can be collected. With regards , however, the use of the DG measure

Finally, R3 also highlights **my interdisciplinary background** as a potential risk for the project. However, contrary to R3, I believe this is also a strength. Firstly, I am not invested in any particular method or outcome, and 'depth' might occasionally come at the cost of a tunnel vision. Given my interdisciplinary approach and network, I believe this puts me in a unique position to make real progress in this area (as acknowledged by R1-2).

Conclusion

In conclusion, all 3 R's recognised the existing knowledge gap and. R1 and R2 offered suggestions, forming logical extensions of the project. While R3 expressed some doubts with regards to the methodology these concern **specific** studies, and **do not affect the set as a whole**.