

Analysis Plan

Our analysis will draw on both parametric and non-parametric methods, supplemented, whenever possible, with accessible graphs including for sub-populations of interest.

Our design requires participants to respond to questions concerning the subjective wellbeing of the two people (A and B) in the various hypothetical scenarios. As described in the Experimental Design section of the pre-trial registration, these are categorized into Low-, Medium- or High-exploitation scenarios (L, M, H), based on data in Ferguson et al. (2026).

1. The first primary outcome is the participants' assessments of the wellbeing of the people in the different scenarios. For each scenario, participants respond on a 5-point Likert scale for each category of subjective assessment (Satisfaction, Happiness, Flourishing) for each of the two people (A and B). Our scenarios are written so that, if there is any exploitation, A is the exploiter and B is the exploited. The question we ask is:

To what extent does this scenario impact person A?

Which is followed by the response scales as seen in the screenshot here:

	falls significantly	falls somewhat	no impact	rises somewhat	rises significantly
A's satisfaction	☐	☐	☐	☐	☐
A's flourishing	☐	☐	☐	☐	☐
A's happiness	☐	☐	☐	☐	☐

The same is then presented on the next page, for person B.

Throughout the analysis we pool responses for each of the wellbeing measures (Satisfaction, Flourishing, Happiness) into the three exploitation levels L, M, H.

We will first present unconditional data, pooling across the two samples.

Our primary analysis will assign numeric values $\{-2, -1, 0, 1, 2\}$ to each point of the Likert response scale, respectively, in the order shown in the screenshot above. For each exploitation level (L, M, H) we will take individual-average responses of the three wellbeing measures for each of A and B, resulting in two composite wellbeing measures: one for A, one for B.¹ We will plot these for the different exploitation levels (L, M, H) with confidence intervals.

Our main hypothesis here is that the reported wellbeing of B is lower in H versus L. We will test for this using a paired t-test.

¹ For those responding to six scenarios (necessarily in the philosopher sample) we will average their responses within each exploitation category (L, M, H) so that every participant contributes once to each of these variables' means.

In secondary analysis, we will look for a richer relationship between exploitation and wellbeing by considering 6 pairwise t-tests: {L vs M, M vs H, L vs H} for each of {A, B}. We will report both raw confidence intervals and p-values as well as p-values adjusted for multiple testing.

We will also plot these data and perform both the primary and secondary analysis above for each sample.

We will test for differences in the aggregated wellbeing measure across the samples. We will do that with 6 pairwise t-tests for each of {L, M, H} x {A, B}. We will report both raw confidence intervals and p-values and p-values adjusted for multiple testing.

In exploratory analysis, we will perform regression analyses. This will feature the averaged measure of wellbeing per actor as the dependent variable, but also the three separate measures, for each of A and B. We will include (i) dummies for the “treatments” of exploitation levels M and H, so that L is the reference category; (ii) a dummy for which sample the respondent belongs to; and (iii) controls such as demographics. We will also include interactions between (i) & (ii), and (i) & (iii), to detect any heterogeneous “treatment” effects. We will additionally include exploratory regression analyses of the philosopher sample to report any potential effects of control measures exclusive to that sample.

In supplementary analysis, we will perform a battery of 18 tests each similar to the primary analysis but with the disaggregated wellbeing measures, i.e., {L vs M, M vs H, L vs H} x {A, B} x {Satisfaction, Happiness, Flourishing}. We will report both raw confidence intervals and p-values and p-values adjusted for multiple testing.

We will also provide an extension which revisits the analyses above with only an ordinal interpretation of the Likert response scales (in place of the earlier cardinal interpretation). We will look at a coarsening of the data in the form of four categories. Specifically, each participant’s responses for each actor in each scenario will fall into one of the following patterns in terms of their three ratings, one for each of Satisfaction, Happiness and Flourishing:

1. at least one positive change and no negative;
2. all three ratings neutral;
3. at least one negative and no positive;
4. a mixture of positive and negative ratings.

We will report the proportion of participants with responses within one of these four categories for each of A and B and potentially also for different combinations.

We will also provide descriptive principal component analyses (separately for A and B) of the wellbeing measures.

2. The second primary outcome is participants’ assessment of how wrong A’s behavior is. For each scenario, after the judgments about A’s and B’s wellbeing, participants answer:

How wrong is A’s behavior towards B?

This is followed by the 5-point Likert (not at all, a little, moderately, significantly, extremely).

The analyses proceed similarly to the wellbeing analysis.

We will first present unconditional data pooling across the two samples.

Our primary analysis will assign numeric values {0, 1, 2, 3, 4} to each point of the Likert response scale, respectively, in the order shown in the screenshot above. We will report and plot the mean response for each exploitation level (L, M, H) with confidence intervals.²

Our main hypothesis here is that mean reported wrongfulness is higher in H than L. We will test for this using a paired t-test.

In secondary analysis, we will look for a richer relationship between exploitation and wrongfulness with three t-tests: {L vs M, M vs H, L vs H}. We will report both raw confidence intervals and p-values and p-values adjusted for multiple testing.

We will also plot these data and perform both the primary and secondary analysis above for each sample.

We will test for differences in the wrongfulness measure across the samples. We will do that with 3 pairwise t-tests for each of {L, M, H}. We will report both raw confidence intervals and p-values and p-values adjusted for multiple testing.

In exploratory analysis, we will perform regression analyses. This will feature the wrongfulness response as the dependent variable and otherwise follow the specification given for the wellbeing analysis above.

As supplementary analysis, we will include exploratory regression analyses of the philosopher sample to report any potential effects of control measures exclusive to that sample.

3. The third primary outcome concerns any norm participants have around wrongfulness. For each scenario, after the wrongfulness judgment, participants answer the following:

Which do you think is the most popular answer to the last question (How wrong is A's behavior towards B?) among all other academic philosopher respondents?

or, if part of the Prolific sample:

Which do you think is the most popular answer to the last question (How wrong is A's behavior towards B?) among all other Prolific respondents?

(The question carries the incentives as detailed in the Experimental Design.) This is followed by the same 5-point Likert scale as the wrongfulness question (not at all, a little, moderately, significantly, extremely).

In this case, the focus of our **primary analysis** will be non-parametric comparison of the two wrongfulness distributions within categories.

² For those responding to six scenarios (necessarily in the philosopher sample) we will average their responses within each exploitation category (L, M, H) so that every participant contributes once to each of these variables' means.

Secondary analysis will use regression methods to explore the correlates of norm perception at the individual-vignette level.

References

Ferguson, B., Matthews, P. H., Ronayne, D., & Veneziani, R. (2026). "What exploitation is"
American Journal of Political Science