Title

A meta-analysis of the effects of experimentally-manipulated demand characteristics

Authors

[redacted]

Research Questions

Overall

What is the overall effect of explicit demand characteristic manipulations on their referenced outcome(s)-of-interest?

To what extent are the effects of explicit demand characteristics manipulations consistent across studies?

Study Characteristic Moderators

We will examine if the effect of explicit demand characteristics manipulations are moderated by any of the following variables:

- Whether participants are students (vs. online workers on platforms like MTurk)
- Whether participants are paid (vs. volunteers)
- Whether the study was conducted in-person (vs. online)
- Whether the effect size compares positive demand characteristics information, negative demand characteristics information, nil demand characteristics information, and/or no demand characteristics information (i.e., a control group)
- Whether the effect size comparison is between- or within-subjects
- Whether the study is published

Conceptual Moderators

If we can obtain the necessary resources, we also intend to examine if the effects of demand characteristics manipulations are moderated by post-hoc ratings of:

- Motivation: Perceived motivation to confirm the experimenter's hypothesis
- Opportunity: Perceived opportunity to control the outcome-of-interest

Hypotheses

This work is mostly exploratory. However, for the conceptual moderators, we are interested in testing the following hypotheses:

H1: The effect of demand characteristic manipulations will be moderated by motivation and opportunity (as predicted by Coles et al. 2022; Rosnow and Rosenthal, 1997)

H2: For the effect of demand characteristics manipulations, there will be a two-way interaction between motivation and opportunity. The simple effect of motivation will be larger when opportunity is high vs. low (as predicted by Coles et al. 2022; Rosnow and Rosenthal, 1997)

H3: When participants are not motivated to confirm or disconfirm the experimenter's hypothesis, the effect of demand characteristics manipulations will be:

- H3.0: Zero (as predicted by Rosnow and Rosenthal, 1997)
- H3.1: Positive and non-zero (as predicted by Coles et al. 2022)

H4: When participants do not have the opportunity to adjust their responses, the effect of demand characteristic manipulations will be:

- H4.1: Zero (as predicted by Rosnow and Rosenthal, 1997)
- H4.2: Positive and non-zero (as predicted by Coles et al. 2022)

Data collection procedures

Our literature search strategy was developed in consultation with a librarian at [redacted]. To identify relevant articles, we searched APA PsycInfo on January 12, 2022 using relatively broad search terms: "demand characteristics" OR "hypothesis awareness". This yielded 850 records.

We also released calls for unpublished studies on the Society for Personality and Social Psychology Open Forum; [redacted] personal Twitter; and the Facebook Psychological Methods Discussion Group and PsychMAP group. This yielded 3 additional records.

Inclusion criteria

To be included in the meta-analysis, records must have met the following inclusion criteria:

- The record documented primary research (as opposed to a review of other findings)
- The researcher manipulated what participants were told about the hypothesized relationship between an independent and dependent variable. (See below for minor exceptions.)

- Participants' responses on the dependent variable described (or strongly implied) in the demand characteristics manipulation were measured.
- Information necessary for computing effect sizes was included.

In a few scenarios, we excluded observations where the researcher's manipulation of demand characteristics was confounded. For example, Sigall, Aronson, and Van Hoose (1970) told one group of participants that they (a) expected an increase in number copying behavior, but (b) this increase would be indicative of a personality disorder. This condition intentionally confounds hypothesis awareness with evaluation apprehension and was thus excluded.

Explanation of existing data and sample size

Thus far, we have coded 232 effect sizes from 34 studies. However, this number will likely change as we discover additional relevant records. (E.g., librarians are searching for several records that we indicated as potentially relevant during the record screening stage, but had difficulty accessing; Reviewers may also recommend additional searches.)

To pilot data processing and analysis procedures, a small subset of the data were analyzed at an earlier stage in the project. At this stage, we found preliminary evidence of a medium-sized, highly-heterogeneous, positive effect of demand characteristics. However, we have not yet examined the full dataset or tested moderators or hypotheses-of-interest.

Measured variables

Effect Size: Cohen's standardized *d*

For studies with between-subject designs and continuous measures:

- *M* and *SD* reported
    - Formula: Cooper, Hedges, & Valentine, 2009; p. 226.
- *t*-values reported
    - Formula: Cooper, Hedges, & Valentine, 2009; p. 228
- *F*-values reported
    - Formula: Cooper, Hedges, & Valentine, 2009; p. 228
- *p*-values reported
    - Formula: Cooper, Hedges, & Valentine, 2009; p. 228

For studies with between-subject designs and categorical measures:

- Formulas:

    - Borenstein et al. 2011; p. 36; Equation 5.8
    - Borenstein et al. 2011; p. 36; Equation 5.9
    - Borenstein et al. 2011; p. 47; Equation 7.1

For studies with within-subject designs:

- *M* and *SD* reported
    - Formulas:
        - Cooper, Hedges, & Valentine, 2009; p. 229
        - http://handbook.cochrane.org/chapter_16/16_4_6_1_mean_differences.htm
- *t*-values reported
    - Formula: Cooper, Hedges, & Valentine, 2009, p. 229
- *F*-values reported:
    - Formula: Cooper, Hedges, & Valentine, 2009; p. 229

Study Characteristic Moderators

- Whether participants are students (yes, no)
- Whether participants are paid (yes, no)
- Whether the study was conducted online (yes, no)
- Type of demand characteristic comparison
    - Positive demand (participants told about hypothesized effect that amplifies relationship between IV and DV) vs. control (participants not told a hypothesis)
    - Positive demand vs. nil demand (participant told that the researcher does *not* expect a relationship between IV and DV)
    - Positive demand vs. negative demand (participants told about hypothesized effect that reverses relationship between IV and DV)
    - Negative demand vs. control
    - Negative demand vs. nil demand
    - Control vs. nil demand

- Comparison design (between-subjects, within-subjects)
- Whether the study is published (yes, no)

Conceptual Moderators

We aspire to collect *post-hoc* measures of motivation and opportunity through a new set of participants. Here, we describe our *preliminary plans*.

For each demand characteristics condition in each study, a new set of participants will review key study details and report the extent to which they would be motivated to provide hypothesis-consistent responses. These ratings will be made on a eleven-point scale (-5 = Extremely motivated to respond in a manner that is inconsistent with the experimenter's hypothesis; 0 = Not motivated to respond adjust responses based on the experimenter's hypothesis; 5 = Extremely motivated to respond in a manner that is consistent with the experimenter's hypothesis).

For each condition, the motivation score will be computed by averaging participant ratings. For the control conditions, we will assume motivation is 0 because participants were not given information about the experimenter's hypothesis. For each effect size, we will sum the motivation scores for the two conditions being compared. Doing so allows us to accommodate the fact that some studies do not use a control group, instead comparing two demand characteristics conditions.

We will use a similar approach for measuring the extent to which participants can intentionally adjust their responses to each dependent variable included in the meta-analysis (0 = not at all able to change the response to 10 = extremely capable of changing the response)

Statistical Models

Meta-regression with robust variance estimates

For overall effect size estimates and subgroup analyses (i.e., analyses split by potential moderators), we plan to use meta-regression with robust variance estimates (Hedges, Tipton, & Johnson, 2010).

We will use the correlated effects weighting scheme with the default assumed within-study effect-size correlation, r =.80. Nonetheless, for our *overall effect size analysis*, we will perform sensitivity analyses that test the following values: r = 0, .20, .40, .60, 1.

To calculate Cohen's *d* for within-subjects design we need the correlation between the pre- and post- measures. This correlation is rarely reported, so we will assume a correlation of r = .50. Nonetheless, for our *overall effect size analysis,* we will perform sensitivity analyses that test the following values: r = .10, .30, .70, .90.

Subgroup analyses

Publication bias

To estimate, correct, and evaluate the impact of publication bias, we will use three approaches: (1) PET-PEESE analyses (Stanley and Doucouliagos, 2013), (2) three-parameter selection modeling (Vevea & Hedges, 1995), and (3) the Mathur and VanderWeele (2020) sensitivity analyses.

Because PET-PEESE and three-parameter selection modeling assume statistical independence, we will run these analyses with aggregated dependent effect sizes (using the Borenstein, 2009 approach). The default assumed correlation between dependent effect sizes in the Borenstein (2009) aggregation method is .50. We will use this default value to inform our conclusion, but re-run the analyses with the following values: r = .10, .30, .70, .90.

Inference criteria

$p < .05$