

Preregistration for Quantitative Research in Psychology (PRP-QUANT) Template

Title

T1 Title

The title should be focused and descriptive, using relevant key terms to reflect what will be done in the study. Use title case (<https://apastyle.apa.org/style-grammar-guidelines/capitalization/title-case>).

Pretesting Video Speeches to Manipulate Perceptions and Examining Strength of Instruments for an Experiment on Servant Leadership and Performance

T2 Contributors, Affiliations, and Persistent IDs (recommend ORCID iD)

Provide in separate entries the full name of each contributor, each contributor's professional affiliation, and each contributor's persistent ID. See ORCID iD for an example of persistent ID (<https://orcid.org/>). Optional: include the intended contribution of each person listed (e.g. statistical analysis, data collection; see CRediT, <https://casrai.org/credit/>).

Annika F. Schowalter, University of Bamberg, <https://orcid.org/0000-0001-9478-7438>
Judith Volmer, University of Bamberg, <https://orcid.org/0000-0003-4476-6537>

T3 Date of Preregistration

This is assigned by the system upon preregistration submission.

T4 Versioning information

This is assigned by the system upon submission of original and subsequent revisions. Should be a persistent identifier, if not a DOI.

T5 Identifier

This unique identifier is assigned by the system upon submission.

T6 Estimated duration of project

Include best estimate for how long the project will take from preregistration submission to project completion.

The pretest is part of a resubmission of a proposal at *The Leadership Quarterly* (Special Issue Call: *Beyond the ritualized use of questionnaires: Toward a science of actual behaviors and psychological states*).

T7 IRB Status (Institutional Review Board/Independent Ethics Committee/Ethical Review Board/Research Ethics Board)

If the study will include human or animal subjects, provide a brief overview of plans for the treatment of those subjects in accordance with established ethical guidelines. If appropriate institutional approval has been obtained for the study, provide the relevant identifier here. If the study will be exempt from ethical board review, provide reasoning here.

The experiment was approved by the local Ethical Review Board (University of Bamberg).

T8 Conflict of Interest Statement

Identify any real or perceived conflicts of interest with this study execution. For example, any interests or activities that might be seen as influencing the research (e.g., financial interests in a test or procedure, funding by pharmaceutical companies for research).

We have no conflicts of interest to declare.

T9 Keywords

Include terms specific to your topic, methodology, and population. Use natural language and avoid words used in the title or overly general terms. If you need help with keywords, try a keyword search using your proposed keywords in a search engine to check results.

Stewardship, authenticity, instrument strength

T10 Data accessibility statement and planned repository

"We plan to make the data available (yes / no)

If "yes", please specify the planned data availability level by selecting one of the options:

- Data access via download; usage of data for all purposes (public use file)
- Data access via download; usage of data restricted to scientific purposes (scientific use file)
- Data access via download; usage of data has to be agreed and defined on an individual case basis
- Data access via secure data center (no download, usage/analysis only in a secure data center)
- Data available upon email request by member of scientific community
- Other (please specify)

Yes.

Data access via download; usage of data for all purposes (public use file)

T11 Optional: Code availability

We plan to make the code available (yes / no).
If "yes", please specify the planned code availability level (use same descriptors of data in T10).

Yes.
Code available upon email request by member of scientific community

T12 Optional: Standard lab practices

Standard lab practices refer to a (timestamped) document, software package, or similar, which specifies standard pipelines, analytical decisions, etc. which always apply to certain types of research in a lab. Specify here and refer to at the appropriate positions in the remainder of the template:
We plan to make the standard lab practices available (yes / no).
If "yes", please specify the planned standard lab practices availability level (use same descriptors of data in T10).

Yes.
Data and procedure available upon email request by member of scientific community

Abstract

(150 words)

A1 Background

(See introduction I1)

The concept of servant leadership has received growing interest in recent years. To date, however, there are no clean causal investigations of the effects of servant leadership on followers' performance because of endogeneity issues. Moreover, the concept has mostly been studied via questionnaires, which can lead to disputable policy implications.

A2 Objectives and Research questions

(See introduction I2)

We are the first to use an instrumental variable approach in the context of servant leadership and created new material based on the literature. This pre-study thus serves to test how followers perceive the leader manipulation and to examine the strength of our instrumental variables (i.e., manipulation of leadership style and leader gender and agreeableness).

A3 Participants

(See methods M4)

We aim to recruit at least 200 working-age individuals using a common panel provider.

A4 Study method

(See methods M10-14)

Our hypotheses will be tested by means of an experimental online study with a 2 (servant leadership vs. neutral speech) x 2 (male vs. female leader) factorial design.

Introduction

(no word limit)

I1 Theoretical background

Provide a brief overview that justifies the research hypotheses.

Initially introduced by Greenleaf (1977), the concept of servant leadership has received growing interest in recent years (Eva et al., 2019; Gardner et al., 2020). In Gardner et al.'s (2020) review of the third decade of leadership research in *The Leadership Quarterly*, for instance, servant leadership was one of the focal leadership theories among the top 50 frequently cited papers. It can be defined as an "(1) other-oriented approach to leadership (2) manifested through one-on-one prioritizing of follower individual needs and interests, (3) and outward reorienting of their concern for self towards concern for others within the organization and the larger community" (Eva et al., 2019, p. 114). One reason for the increasing popularity of the leadership concept is the positive associations with various outcomes (e.g., individual, team, and organizational performance; community citizenship behaviors; organizational commitment), even beyond other leadership concepts like authentic, ethical, and transformational leadership (e.g., Eva et al., 2019; Hoch et al., 2018; Lee et al., 2020; Parris & Peachey, 2013). To date, however, there are no clean causal investigations of direct and mediating effects of servant leadership on followers' performance (Eva et al., 2019; Lee et al., 2020). Moreover, the concept has mostly been studied via questionnaires, which can lead to disputable policy implications. For instance, questionnaire measures often confound behavior and evaluations or involve the risk that the results obtained are not caused by real effects but, for example, by participants' consistent responding to items (Alvesson, 2020) or other issues of endogeneity (Sajons, 2020). With our study, we aim to shed light on the causal gap by experimentally manipulating servant leadership (i.e., authenticity and stewardship) and investigating its effect on performance. Moreover, we will include followers' perceptions of servant leadership as a mediating variable (see Figure 1).

Our study will add value to the current research in several ways. First, building on Eva et al.'s work (2019), who examined the research on servant leadership between 1998 and 2018, we will conduct a systematic literature review on articles that investigated servant leadership as an explanatory or a mediating variable. Thereby, we will show that the gap of causal examinations as well as the risk of endogeneity in research on servant leadership persists.

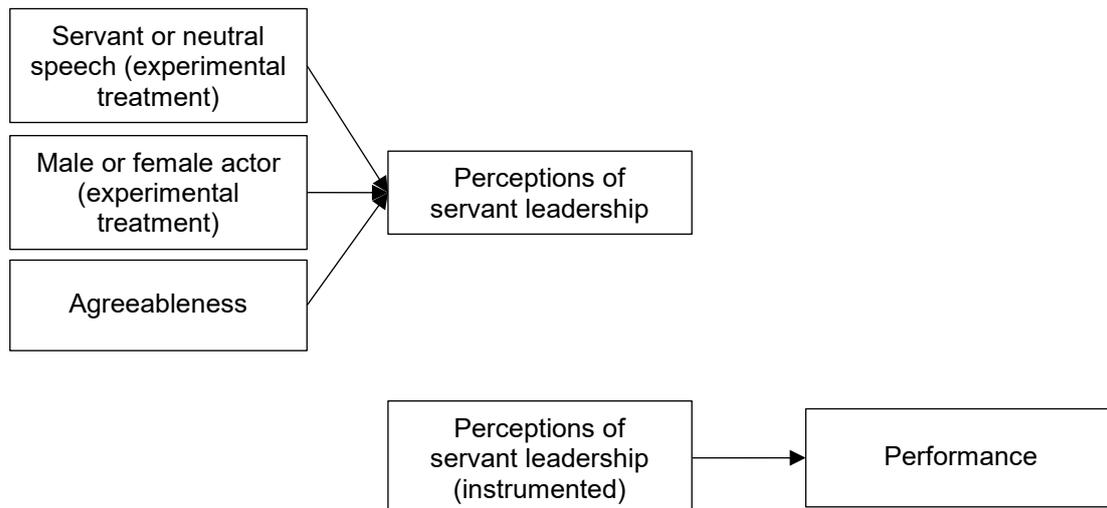
Second, to avoid biases due to endogeneity problems and to make claims regarding causality, experimental studies and/or corrective actions are necessary (Podsakoff & Podsakoff, 2019). Thus, to change the deficient state of affairs in servant leadership research, we will examine the effect of servant leadership on followers' performance with an online experiment. In vignettes, it is mostly necessary to explicitly mention the manipulated variable (in our case leader or leadership). Thereby, critical information about the hypothesis is revealed to participants producing a risk of demand effects potentially rendering the results uninformative for policy (Khademi et al., 2021). If participants then adapt their behavior, it remains unclear whether the reason is actual effects of the described leadership behavior or whether participants behave according to what they think is the behavior expected by the researchers (Khademi et al., 2021; Zizzo, 2010). By manipulating leadership via videos, we do not have to explicitly mention the manipulated variable. Moreover, we will ask participants about their servant leadership perceptions after the performance task. By placing the questionnaire after the outcome, we can ensure

obfuscation of the research question and further reduce the risk for demand effects. As per definition, servant leadership is an interactive leadership style (Eva et al., 2019), not all dimensions are equally manipulable in a non-interactive online experiment. Therefore, we will concentrate on two dimensions of servant leadership which can be examined without interaction between leader and followers (i.e., stewardship and authenticity; Van Dierendonck & Nuijten, 2011) and propose a way to cleanly test the effects of the leadership concept.

Third, we will use an instrumental variable estimation approach as leadership perceptions (our mediator) cannot be directly manipulated and thus, are endogenous (Sajons, 2020). Thereby, we will show how the investigation of servant leadership perceptions on performance can cleanly be done. With the instrumental variable estimation, we can interpret the endogenous mediator causally and show that there is a threat of endogeneity when measuring servant leadership perceptions.

Fourth, we will include a behavioral measure of employees' performance in our experiment. There is often a wide gap between intentions and actual behavior (Orbell & Sheeran, 1998) and questionnaire measures can be severely biased, for instance, by common method variance or demand effects (as questionnaires are seldom consequential and the answers are often easy to adapt to the inferred experimenter's objective), potentially leading to misdirection in science and policy implications (Alvesson, 2020). With our performance task, we can capture real behavior and thus, effectively address these problems. In addition, we will follow Antonakis' (2017) call to design realistic experiments by not embedding the experimental tasks into a fictitious story but framing them (veridically) as an online experiment. Additionally, we will motivate the participants with a real outcome (i.e., donations to a charity) depending on their performance (Meslec et al. 2020). These measures will enhance external validity and simultaneously maintain a high internal validity.

Figure 1
Hypothesized Model



Note. The hypothesized mediation model. The pre-study is used to test the upper part of the model.

As there are no existing manipulations of servant leadership besides written vignettes describing a leader from a third-person point of view, we created new material based on the literature. Moreover, we are the first to use an instrumental variable approach in the context of servant leadership. The pre-study thus serves to test how followers perceive the

leader manipulation and to examine the strength of our instrumental variables (i.e., manipulation of leadership style and leader gender and agreeableness). Furthermore, with the help of the results of the pre-test, we can calculate an informed power analysis for the experiment.

Servant Leadership

Servant leadership is described as a holistic leadership approach combining both task and people focus by being people-centered without ignoring performance expectations (Eva et al., 2019; Van Dierendonck & Nuijten, 2011). In their review, Eva et al. (2019) recommended three measures of servant leadership; one of them is the Servant Leadership Survey by Van Dierendonck and Nuijten (2011). The latter identified eight dimensions of servant leadership: (1) Empowerment aims at motivating employees to be proactive and self-confident and focusses on their personal development; (2) accountability means that people know their boundaries and what is expected from them, and are held responsible for their performance; (3) standing back refers to the extent to which a leader acknowledges others' contributions and shares credits for successfully accomplished tasks; (4) humility describes the leader's ability to put their own accomplishments and capacities in an appropriate perspective, knowing about own strengths and limitations, and seeking others' contributions to overcome the latter; (5) authenticity is the degree to which a leader is true to themselves and shows their true emotions and motivations to the employees; (6) courage describes the attitude to question old approaches and to try out new ones, and to willingly face challenges, in accordance with own beliefs and values; (7) forgiveness means that one accepts others despite mistakes and wrongdoings and is not eager to revenge or get even; (8) stewardship is about acting as a role model and taking responsibility for the organization beyond genuine self-interest.

Amongst others, research indicates a positive relationship between servant leadership and organizational citizenship behavior (OCB; e.g., Liden et al., 2008), proactive behavior (e.g., Bande et al., 2016), psychological well-being (e.g., Gotsis & Grimani, 2016), collaboration (e.g., Garber et al., 2009), and individual (e.g., Liden et al., 2008), team (e.g., Sousa & Van Dierendonck, 2017), and organizational performance (e.g., Choudhary et al., 2013). Servant leadership shows medium to large correlations with transformational leadership ($r = .32-.79$; Eva et al., 2019), which consists of the four aspects idealized influence, inspirational motivation, intellectual stimulation, and individual consideration (Bass, 1999). However, there are some differences between the two leadership styles. Servant leaders' motivation lies (different from traditional leadership styles) beyond self-orientation, beyond a concentration on their own advancement (Eva et al., 2019).

Transformational leaders predominantly focus on organizational goals whereas servant leaders put their employees' needs before the organizational goals (Van Dierendonck et al., 2014). Additionally, there exists evidence that servant is associated with performance outcomes above and beyond transformational leadership (Eva et al., 2019; Hoch et al., 2018). Hoch et al. (2018) found, for instance, in their meta-analysis that servant leadership added about 12% to the explained variance in outcomes beyond transformational leadership.

In our experiment, we will concentrate on the two dimensions stewardship and authenticity as the other dimensions cannot be manipulated cleanly within a video due to their interactive elements: The dimension of empowerment requires that leaders know their individual followers' skills and abilities to be able to support their personal development (Van Dierendonck & Nuijten, 2011). An essential part of humility is dealing with criticism by others (Van Dierendonck & Nuijten, 2011) which would make it necessary that others criticize the leader. Given that there is no real criticism, the leader could tell a story about a past situation or how they would handle such situations. This would not depict real behavior but resemble an experimental vignette. As mentioned before, vignettes are not very informative for policy, which is why we decided against manipulating humility in the

video. A similar issue applies to courage as both items in the Servant Leadership Survey (Van Dierendonck & Nuijten, 2011) concern taking risks, one of them without being sure of the support by one's supervisor. As with humility, risky behavior in the sense of servant leadership (in line with the leader's values and beliefs) and the leader defending their own opinion even if the own supervisor does not necessarily support it, cannot cleanly or realistically be manipulated in the video speech. This makes it hardly possible to thoroughly examine if leader behavior and followers' perceptions (i.e., our mediator) coincide. Likewise, accountability, standing back, and forgiveness require preceding follower behavior in terms of performance or failures (Van Dierendonck & Nuijten, 2011). Therefore, our video manipulation as well as our explanations regarding the associations between servant leadership, its perceptions, and performance will focus on authenticity and stewardship. As the two dimensions are the only ones cleanly testable with our video manipulation, we will still refer to their combination as servant leadership in our article. Authenticity and creating value for the community (which is part of stewardship) are also two essential dimensions Anderson and Sun (2017) extracted when they synthesized different measures for servant leadership.

Stewardship

Stewardship includes a "focus on a long-term perspective that takes into account all stakeholders" (Eva et al., 2019, p. 113). The term stakeholders subsumes a wide variety of individuals and groups, from one's followers to the organization to the general social world, also referred to as the common good (Eva et al., 2019; Pircher Verdorfer & Peus, 2014; Van Dierendonck & Nuijten, 2011). Additionally, knowing where to take their organization and providing a long-term vision, servant leaders give their followers' direction (Van Dierendonck & Nuijten, 2011).

Authenticity and Authentic Leadership

Authenticity, as defined by Van Dierendonck and Nuijten (2011) includes "being true to oneself, accurately representing – privately and publicly – internal states, intentions, and commitments (Peterson and Seligman 2004)" (Van Dierendonck & Nuijten, 2011, p. 252). According to Eva et al. (2019) authentic and servant leadership share that leaders should be authentic and true in their interactions with others. In recent years, several leadership scholars started to discuss whether the current positive leadership constructs (e.g., servant leadership, transformational leadership, ethical leadership, or authentic leadership) can be differentiated or rather measure an overarching positive leadership construct (e.g., Hoch et al., 2018). As mentioned, primary studies, as well as meta-analyses, found that servant leadership was associated with performance outcomes over and above transformational leadership. Even though the contributions of authentic leadership were less than half as high (Hoch et al., 2018), the high correlations between servant and authentic leadership as well as the use of similar items to some extent corroborate the thesis of overlap between them (Lemoine et al., 2019; see also Banks et al., 2018). Authentic leadership is defined "as a process that draws from both positive psychological capacities and a highly developed organizational context, which results in both greater self-awareness and self-regulated positive behaviors on the part of leaders and associates, fostering positive self-development" (Luthans & Avolio, 2003, p. 243). There are different theoretical views on the relationship between servant and authentic leadership (e.g., Eva et al., 2019; Lemoine et al., 2019). For instance, Eva et al. (2019) point to a different underlying mindset stating that "servant leaders are authentic not for the sake of being authentic, but because they are driven either by a sense of higher calling or inner conviction to serve and make a positive difference for others" (p. 113) which is not part of the authentic leadership framework. Alvesson and Einola (2019) call into question if authentic leadership can be (different than servant leadership) categorized as moral leadership at all.

At least the subdimension of authentic leadership *relational transparency* overlaps with the

authenticity dimension of servant leadership. It can be defined as “showing one's true self to others and openly, but appropriately, sharing information regarding one's true thoughts and emotions” (Banks et al., 2016). Therefore, as our experiment could also be informative for authentic leadership research, we plan to include a systematic review to show that this leadership concept as well has not been cleanly causally examined so far.

Eva et al. (2019) give suggestions on how to manipulate authenticity as a servant leader behavior, for instance: “The leader follows through on their actions (accountability, integrity and honesty) and is humble in the language they use” as an actor-led behavior (p. 126). As Alvesson and Einola (2019) explain, these behaviors can neither be manipulated nor assessed by followers. Taking a closer look at the dimension of authenticity in Van Dierendonck and Nuijten's (2011) Servant Leadership Survey, we find that they do not include authenticity in the sense of being honest and of integrity. They rather conceptualize it as being sensitive and showing one's emotions and shortcomings. This enables us to experimentally manipulate authenticity. Thus, we rely on the operationalization of authenticity in the Servant Leadership Survey (Van Dierendonck & Nuijten, 2011) for creating the videos.

Servant Leadership and Performance

Researchers propose several ways of how servant leadership could enhance performance. As cited in the introduction, Eva et al. (2019) developed a three-part definition of servant leadership. The third part corresponds to the dimension of stewardship: Servant leaders are concerned for others' well-being, ranging from the individual employee to the larger community. This includes that leaders promote the growth of followers and other resources within the organization, which changes, as the authors postulate, followers' orientation from egoistical to other-serving. This, in turn, should enable followers to be productive and to make a difference to the social world (Eva et al., 2019). Theories used to explain the effects of servant leadership are, for example, social exchange theory (Blau, 1964), social learning theory (Bandura, 1977), or social identity theory (Tajfel, 1978). Although social exchange theory is often used as a theoretical foundation in servant leadership research, the other two theories are suggested to be more appropriate to capture the long-term transforming effect of the concept (Eva et al., 2019). According to social learning theory, if followers perceive their servant leaders as credible role models (i.e., because of their altruistic and other-serving motivation and behavior), they observe their leaders and strive for similar attitudes, behaviors, and values (Eva et al., 2019). As stated by social identity theory, followers' identification with their group is central to change their behavior. As servant leaders focus on their followers and are authentic, they can create strong relationships with their followers and thereby foster their followers' identification with their group. In that way, employees also behave more beneficial to the organization (Chen et al., 2015). Thus, stewardship and authenticity are two crucial dimensions, enabling the positive effects of servant leadership on performance.

The Mediating Effect of Leadership Perceptions

As mentioned before, the effect of servant leadership on various outcomes has mostly been measured via questionnaires. Questionnaires, however, cannot measure real behavior but only followers' perceptions. These perceptions are not only influenced by leader behavior but also by other (omitted) variables like followers' individual differences (Felfe & Heinitz, 2010). Hence, perceptions are endogenous and, if not properly accounted for, prevent from drawing solid causal conclusions (Sajons, 2020). For instance, in field studies using questionnaires, it remains unclear whether different ratings among followers can be ascribed to differential leader behaviors or rater traits (Bono et al., 2012). As an example for a rater trait, it could be the case that followers scoring high on agreeableness rate their leader more leniently (Bernardin et al., 2000). At the same time, the outcome may be influenced by followers' agreeableness: In an experimental study, agreeable individuals may work harder wanting to support the researchers. As a result, correlations

between leader behavior and outcomes may be biased (Hansbrough et al., 2015). As well, simultaneity bias (i.e., mutual influence of the explanatory variable and the dependent variable) cannot be excluded (Güntner et al., 2020). Thus, as perceptions are not directly manipulable (Sajons, 2020), corrective actions are necessary to build solid knowledge about how servant leadership affects performance.

To contribute to the understanding of the potential influence of servant leadership on task performance, mediated by the perception of servant leadership, we plan to study the effect using three instrumental variables (i.e., leadership style, and leader gender, and agreeableness). Using more than one instrumental variable can reduce standard errors (Semadeni et al., 2014) and allows for testing the exclusion restriction (Sajons, 2020).

Thus, we will follow Antonakis et al.'s (2010) recommendation and use two additional instruments besides the manipulation of leadership: another manipulated variable (i.e., leader's gender) and a personality variable (i.e., followers' agreeableness).

Instrumental variables have to be relevant, exogenous and fulfill the exclusion restriction (Sajons, 2020). As we plan to directly manipulate authenticity and stewardship through our videos (servant leadership vs. neutral), the randomly assigned leadership manipulation is exogenous and a theoretical cause for servant leadership perceptions. We will test the relevance condition also statistically in this pre-test.

Regarding the exclusion restriction, one could assume other channels than servant leadership perceptions through which our leadership manipulation influences the outcome, for instance, comprehensibility of the content or other leadership styles. However, the two videos (neutral and servant) only differ in the content of the speech and the nonverbal behavior (i.e., expression of emotions) regarding the two servant leadership dimensions stewardship and authenticity. The instructions for the task are additionally presented in a written format so that there should be no differences between the groups regarding the understanding of the performance task. Therefore, the instrumented perceptions of servant leadership should be the only channel through which our manipulation influences performance, and the exclusion restriction should be fulfilled. Satisfying the three conditions, we conclude that servant leadership versus the neutral condition should be an appropriate instrumental variable.

Previous research has shown that leaders are rated differently depending on their gender. For instance, Butler and Geis (1990) found that women were rated less favorably than men although they had shown the same behavior. One explanation for this bias is that leadership is stereotypically seen as a male domain and therefore the fit of women in leadership positions may be perceived to be lower (Hansbrough et al., 2015). However, it is still unclear, how exactly the perceptions of servant leadership vary depending on leader gender (Lehrke & Sowden, 2017; Lemoine & Blum, 2021). Because of the communal aspects, servant leadership could match stereotypes about women (Lemoine & Blum, 2021) and lead to more positive perceptions of a female leader. As we manipulate leader's gender in our videos, this instrument is per definition exogenous.

Agreeableness has shown to be correlated with followers' perceptions of leadership behavior (Schyns & Sanders, 2007; Wang et al., 2019), but, different than the other Big Five personality traits, it is hardly or not directly related to cognitive performance (Demetriou et al., 2003; Furnham et al., 2005; O'Connor & Pounonen, 2007; Sutin et al., 2021). These findings indicate that agreeableness should influence performance only indirectly via servant leadership perceptions. As personality variables are often suggested as instrumental variables (e.g., Antonakis et al., 2010), we will consider agreeableness as a potential instrumental variable despite our concerns regarding the exclusion restriction (see our explanation regarding agreeableness as an omitted variable).

Thus, we will examine the relevancy of agreeableness and leader's gender as instrumental variables. If relevancy is given, we will test the exclusion restriction in our main experiment. Given that the exclusion restriction would not be fulfilled if the leader's gender or followers' agreeableness is incorporated in the model, we will reject the

respective variable as an instrument and conduct our analyses only with our leadership manipulation and as appropriate with the remaining second instrument.

I2 Objectives and Research question(s)

Outline objectives and research questions that inform the methodology and analyses (below).

With this study, we aim to pre-test our manipulation and instrumental variables. Moreover, we derive estimates for the power analysis to determine the appropriate sample size for the planned experiment.

Research questions:

1. Do individuals in the servant leadership condition perceive the leader as more servant than individuals in the neutral condition?
2. Are the instrumental variables strong enough for the planned analyses?

I3 Hypothesis (H1, H2, ...)

Provide hypothesis for predicted results. If multiple hypotheses, uniquely number them (e.g., H1, H2a, H2b,) and refer to them the same way at other points in the registration document and in the manuscript.

H1: Individuals exposed to the servant speech (i.e., stewardship and authenticity) will perceive the leader as more servant than individuals exposed to the standard speech.

H2: The manipulation of the leadership style (servant leadership vs. neutral) is a relevant instrumental variable to instrument followers' perceptions of servant leadership.

H3: The manipulation of the actor (male vs. female) is a relevant instrumental variable to instrument followers' perceptions of servant leadership.

H4: Participants' agreeableness is a relevant instrumental variable to instrument followers' perceptions of leadership on task performance.

I4 Exploratory research questions (if applicable; E1, E2,)

If planning exploratory analyses, provide rationale for them here. If multiple exploratory analyses, uniquely number them (E1, E2, ...) and refer to them in the same way in the registration document and in future publications.

We plan to investigate whether our leadership manipulation also influences perceptions of charismatic leadership. If there are no differences in perceptions of charismatic leadership between the two conditions (servant leadership vs. neutral), this would additionally strengthen our argument regarding the exclusion restriction.

Furthermore, we will collect data on the perception of the non-manipulated dimensions of servant leadership (i.e., empowerment, accountability, standing back, humility, courage,

forgiveness) for further tests in case the manipulation does not work as intended.

Method

M1 Time point of registration

Select one of the options:

- Registration prior to creation of data
- Registration prior to any human observation of the data
- Registration prior to accessing the data
- Registration prior to analysis of the data
- Other (please specify; might include if T1 longitudinal data has been analyzed, but T2 has not yet been analyzed)

Registration prior to creation of data

M2 Proposal: Use of pre-existing data (re-analysis or secondary data analysis)

Will pre-existing data be used in the planned study? If yes, indicate if the data were previously published and specify the source of the data (e.g., DOI or APA style reference of original publication). Specify your level of knowledge of the data (e.g., descriptive statistics from previous publications), whether or not this is relevant for the hypotheses of the present study, and how it is assured that you are unaware of results or statistical patterns in the data of relevance to the present hypotheses.

-

Sampling Procedure and Data Collection

M3 Sample size, power and precision

(1) Relevant sample sizes: e.g., single groups, multiple groups, and sample sizes (or sample ranges) found at each level of multilevel data. (2) Provide power analysis (e.g. power curves) for fixed-N designs. For sequential designs, indicate your 'stopping rule' such as the points at which you intend to be viewing your data and in any way analyzing them (e.g., t-tests and correlations, but even descriptively such as with histograms).

As there are no previous data we could rely on to derive a sound estimate of the needed sample size through power analysis, and because the preliminary study is intended to serve as a basis to calculate the sample size for the planned experiment, we follow recommendations in the literature. In total, we aim to recruit at least 200 individuals in accordance with Lonati et al. (2018) who recommend at least 50 individuals per condition in experimental studies. This was also the sample size, Sajons (2020) used for her experiment with two experimentally randomized instrumental variables.

M4 Participant recruitment, selection, and compensation

Indicate (a) methods of recruitment (e.g., subject pool advertisement, community events, crowdsourcing platforms, snowball sampling); (b) selection and inclusion/exclusion criteria (e.g., age, visual acuity, language facility); (c) details of any stratification sampling used; (d) planned participant characteristics (gender, race/ethnicity, sexual orientation and gender identity, SES, education level, age, disability or health status, geographic location); (e) compensation amount and method (e.g., same payment to all, pay based on performance, lottery).

We plan to use a common panel provider, ideally Respondi, with a UK sample. All participants should be English-speaking. To obtain a sample that is as representative as possible, we plan to use quotas reproducing the distribution of age and gender in the working-age population in the UK. The numbers are derived from the latest Census. Using the same sample characteristics in the pretest as in the planned experiment, we want the results to be as comparable as possible to the experiment (where the sample should be representative).

Males: $N = 98$

Females: $N = 102$

Age 18-24 years: $N = 30$

Age 25-29 years: $N = 22$

Age 30-39 years: $N = 43$

Age 40-49 years: $N = 47$

Age 50-59 years: $N = 39$

Age 60-64 years: $N = 19$

Participants will receive the usual compensation from the panel provider. Additionally, the research team will donate 0.03 GBP to the charity World Vision (a non-profit relief, development, and advocacy organization) for each letter that is correctly decoded in the course of the performance task.

M5 How will participant drop-out be handled?

Indicate any special treatment for participants who drop out (e.g., there is follow-up in a manner different from the main sample, last value carried forward) or whether participants are replaced.

We will use only one measurement point.

M6 Masking of participants and researchers

Indicate all forms of masking and/or allocation concealment (e.g., administrators, data collectors, raters, confederates are unaware of the condition to which participants were assigned).

It is an online experiment with a randomized assignment to groups and a manipulation via recorded videos, all researchers will be unaware of the condition to which participants will be assigned.

M7 Data cleaning and screening

Indicate all steps related to data quality control, e.g., outlier treatment, identification of missing data, checks for normality, etc.

Cases with incorrect answers on all three attention check items will be excluded. We will plot the data (histograms and box-whisker diagrams) and check the distribution of the servant leadership, the charismatic leadership, and agreeableness scores as well as the descriptive data for plausibility (such as mean, *SD*, min, max).

We will calculate the reliability and the mean of the servant leadership, the charismatic leadership, and the agreeableness items and check the values for normality.

M8 How will missing data be handled?

Indicate any procedures that will be applied during the analysis to deal with missing data, such as (a) case deletions; (b) averaging across scale items (to handle missing items for some); (c) test of missingness (MAR, MCAR, MNAR assumptions); (d) imputation procedures (FIML vs. MI); (e) Intention to treat analysis and per protocol analysis (as appropriate).

Cases with no data on the variables will be deleted.

M9 Other information (optional)

For example, training of raters/participants or anything else not yet specified.

Conditions and design

M10 Type of study and study design

Indicate the type of study (e.g., experimental, observational, cross-sectional vs. longitudinal, single case, clinical trial) and planned study design (e.g., between vs. within subjects, factorial, repeated measures, etc.), number of factors and factor levels, etc..

We will use an experimental online study with a 2 (servant leadership vs. neutral speech) x 2 (male vs. female leader) factorial design, resulting in four conditions.

M11 Randomization of participants and/or experimental materials

If applicable, describe how participants are assigned to conditions or treatments, how stimuli are assigned to conditions, and how presentation of tests, trials, etc. is randomized. Indicate the randomization technique and whether constraints were applied (pseudo-randomization). Indicate any type of balancing across participants (e.g., assignments of responses to hands, etc.).

Participants will be randomly assigned to groups by a randomization trigger that is linked with filters. The filters allow for showing only one video to each participant. The items to measure servant leadership, charismatic leadership, and agreeableness as well as the sequence of the lines of codes (performance task) will be randomized.

M12 Measured variables, manipulated variables, covariates

This section shall be used to unambiguously clarify which variables are used to operationalize the hypotheses specified above (item I3). Please (a) list all measured variables, and (b) explicitly state the functional role of each variable (i.e., independent variable, dependent variable, covariate, mediator, moderator). It is important to (c) specify for each hypothesis how it is operationalized, i.e., which variables will be used to test the respective hypothesis and how the hypothesis will be operationally defined in terms of these variables. The description here shall be consistent with the statistical analysis plans specified under AP6 (below).

Manipulated variables (four videos in total):

- Leadership style
 - o The manipulation consists of two factors: servant leadership vs. neutral
 - o The variable serves as the independent variable for Hypothesis 1.

Additionally, we will examine its strength as an instrumental variable (Hypothesis 2).

- Actor (leader gender)
 - o The manipulation consists of two factors: male vs. female
 - o As with leadership style, we will test the strength of leader gender as an instrumental variable (Hypothesis 3).

Measured variables:

- Perceptions of servant leadership
 - o The variable (only the dimensions stewardship and authenticity) serves as an outcome to test the effect of the leadership speeches (Hypothesis 1) and is used as the dependent variable in the first of the two-stage least squares (2SLS) procedure to test the relevance of the instrumental variables.
 - o The remaining dimensions serve as a backup.
- Perceptions of charismatic leadership
 - o We will test whether differences in charismatic leadership perceptions occur depending on the leadership manipulation.
- Task performance
 - o Decoding task
 - o We use a short form (only seven minutes) of the experimental task to avoid deception.
 - o We will not use it to test our hypotheses.
- Agreeableness
 - o We will test the strength of agreeableness as an instrumental variable to instrument servant leadership perceptions.

Demographic & control variables

- Participants' age and gender, education
 - o Participants' age and gender are necessary for quota sampling to approach representativeness of the sample regarding the UK working-age population (18-64 years).
 - o We will use the three variables for randomization checks and include them in our analyses as control variables given that there is no full randomization.
- Additionally, we will include one instructed response item (cf. Meade & Craig, 2012) and two questions regarding the content of the study and leadership video as an attention check. If the three questions are answered incorrectly, we will exclude the respective participant.

Operationalization of hypotheses

- Hypothesis 1: β_1 should be significant when servant leadership perceptions are regressed on leadership style.
- Hypothesis 2, 3, and 4: The F -statistic of leadership style exceeds the critical values by Stock and Yogo (2005) for the relevance of the instrument.

M13 Study Materials

Please describe any relevant study materials. This could include, for example, stimulus materials used for experiments, questionnaires used for rating studies, training protocols for intervention studies, etc.

Leadership Manipulation

Participants will be randomly assigned to one of four videos (i.e., servant or neutral leader condition, played by a male or female actor) of equivalent length. Videos are recommended as a more appropriate and realistic manipulation than written vignettes (Antonakis, 2017). In the servant leadership speech, we will manipulate the two dimensions of stewardship and authenticity based on the descriptions by Eva et al. (2019), Pircher Verdorfer and Peus (2014), and Van Dierendonck and Nuijten (2011). We concentrated on these two dimensions as these are the only dimensions in Van Dierendonck and Nuijten's (2011) conceptualization of servant leadership that do not necessarily require interactions with the leader which we could not accomplish with video manipulation.

The neutral speech builds on the standard speech by Meslec et al. (2020; based on Antonakis et al., 2015, 2021). Besides adjusting it to fit the different charity, we adapted the speech in the following ways to establish a clearer distinction from the servant leadership dimensions: We removed the repeatedly mentioned references to the importance of helping the children as this resembles the stakeholder focus and the "strong sense of obligation for the common good" (Pircher Verdorfer & Peus, 2014, p. 3) of stewardship. Moreover, we framed the description of the charity less visionary and more abstract to avoid further confounding with stewardship that includes a long-term vision (Van Dierendonck & Nuijten, 2011). For this purpose, we described World Vision's fields of activity more factually and added some information about the organization that is not included in the servant leadership speech but emphasizes the importance of the organization's work. In this way, we can achieve the same length of the speeches and also ensure that the neutral speech is still motivating and not becoming boring due to repetition. Before the two speeches will be recorded by the actors, we will conduct an objective manipulation check. Two independent coders, unaware of the purpose of the study, will code both speeches sentence by sentence for the presence of the two servant leadership dimensions as well as charismatic leadership tactics (Antonakis et al., 2021). Thereby, we can show that the speeches manipulate only servant leadership but not charismatic leadership. We will calculate coders' agreement with Cohen's κ and compare the occurrences of servant leadership and charismatic leadership tactics between the two speeches.

Questionnaires

After the video, participants will rate the four items to measure authenticity and three items to measure stewardship from the Servant Leadership Scale (Van Dierendonck & Nuijten, 2011). We adapted the items to fit the experimental context (i.e., instead of *leader*, we will use *the person in the video*; we deleted *often*). An example item for authenticity is "The person in the video shows his/her true feelings to the participants" and for stewardship "The person in the video emphasizes the societal responsibility of our work". The items will be rated on six-point Likert scales from 1 = *strongly disagree* to 6 = *strongly agree*. Perceptions of charismatic leadership will be measured as in Antonakis et al. (2021) using the idealized influence and inspirational motivation subscales of the Multifactor Leadership Questionnaire (MLQ, Bass & Avolio, 1995). The 12 items will be rated on six-point Likert scales from 1 = *not at all* to 5 = *frequently, if not always*. An example item is "Talks optimistically about the future".

To measure agreeableness, participants will rate 10 items from the 50-item International Personality Item Pool (IPIP; Goldberg, 1999) on a five-point Likert scale from 1 = *strongly disagree* to 5 = *strongly agree*. An example item is "am interested in people".

Regarding the demographic variables, we will ask for participants' age, gender, and education. Additionally, we will add one instructed response item (see Meade & Craig, 2012) and two questions regarding the content of the study and leadership video to make sure that participants have been attentive.

Cognitive Task

As the performance task, we will use the task by Meslec et al. (2020) where participants are asked to decrypt codes into meaningful English phrases. Therefore, participants will be successively presented with 24 lines of codes and five different coding schemes. They will first have to carefully examine the decoding schemes and choose the appropriate one before they start decoding. To decrypt one line of code, up to three different coding schemes can be necessary. For every correctly decrypted letter within the time limit of 7 minutes, we will donate 0.03 GBP to World Vision. A sample code is “Nccrgvgr sbe rkcrevrpr“, with the solution “Appetite for experience“. In the pre-test, the task is necessary to avoid any deception as explaining the significance of the task is part of the speeches.

M14 Study Procedures

Please describe here any relevant information about how the study will be conducted, e.g., the number and timing of measurement time points for longitudinal research, the number of blocks or runs per session of an experiment, laboratory setting, the group size in group testing, the number of training sessions in interventional studies, questionnaire administration for online assessments, etc.

After participants' consent and questions on age and gender (to check quotas), participants will be introduced to their task (i.e., decrypting letters). We will inform them that for each correctly decrypted letter, we will make a real donation of 0.03 GBP to World Vision. Afterward, we will expose participants to one of four videos based on a 2 (servant leadership speech vs. neutral speech) x 2 (female or male leader) design. Subsequently, we will ask participants to answer questions about their servant leadership perceptions, their charismatic leadership perceptions (including one instructed response item), and their agreeableness. After having worked on the task for 7 minutes, participants will be asked about their education and fill in two more attention check items.

M15 Other information (optional)

Analysis plan

(NOTE: If this varies by hypothesis, repeat analysis plan for each)

AP1 Criteria for post-data collection exclusion of participants, if any

Describe all criteria that will lead to the exclusion of a participant's data (e.g. performance criteria, non-responding in physiological measures, incomplete data). Be as specific as possible.

Cases with no data on the variables will be deleted. Cases with three incorrect answers on the three attention check items will be excluded.

AP2 Criteria for post-data collection exclusions on trial level (if applicable)

Describe all criteria that will lead to the exclusion of a trial or item (e.g. statistical outliers, response time criteria). Be as specific as possible.

AP3 Data preprocessing

Describe all data manipulations that are performed in preparation of the main analyses, e.g. calculation of variables or scales, recoding, any data transformations, preprocessing steps for imaging or physiological data (or refer to publicly accessible standard lab procedure, cf. T12).

We will calculate the means of the servant leadership, charismatic leadership, and agreeableness scale.

AP4 Reliability analysis (if applicable)

Specify the type of scale reliability that will be estimated, whether it is internal consistency (e.g. Cronbach's alpha, omega), test-retest reliability, or some other form (e.g., a

confirmatory factor analysis incorporating multiple factors as sources of variance). In a study involving measure development, researchers should specify criteria for removing items from measures a priori (e.g., largest factor loading magnitude, smallest drop in alpha-if-item removed).

We will calculate McDonald's omega for the servant leadership, charismatic leadership, and the agreeableness scale.

AP5 Descriptive statistics

Specify which descriptive statistics will be calculated for which variables. If appropriate, specify which indices of effect size will be used. If descriptive statistics are linked to specific hypotheses, explicitly link the information given here to the respective hypothesis.

We will calculate means and standard deviations of the servant leadership scale, the charismatic leadership scale, agreeableness, and age.

AP6 Statistical models (provide for each hypothesis if varies)

Specify the statistical model (e.g. t test, ANOVA, LMM) that will be used to test each of your hypotheses. Give all necessary information about model specification (e.g., variables, interactions, planned contrasts) and follow-up analyses. Include model selection criteria (e.g., fit indices), corrections for multiple testing, and tests for statistical violations, if applicable. Wherever unclear, describe how effect sizes will be calculated (e.g., for d-values, use the control SD or the pooled SD).

To check the randomized group assignment, we will regress participants' age, gender, and level of education on the manipulations. If randomization was not successful regarding one or more of these variables, we will include the affected variables as control variables.

To test Hypothesis 1, we will use OLS regression as well.

To be relevant, the instruments have to significantly correlate with the servant leadership scale ($Cov(z,x) \neq 0$). H2, H3, and H4 can therefore be accepted if the first-stage F -statistics of the two-stage least squares (2SLS) procedure exceed the critical values by Stock and Yogo (2005).

AP7 Inference criteria

Specify the criteria used for inferences (e.g., p values, Bayes factors, effect size measures) and the thresholds for accepting or rejecting your hypotheses. If possible, define a smallest effect size of interest. If inference criteria differ between hypotheses, specify separately for each hypothesis and respective statistical model by explicitly referring to the numbers of the hypotheses. Describe which effect size measures will be

reported and how they are calculated.

Regarding the regression analysis (H1), the p -Value of the coefficient should be below the significance level of $\alpha = .05$ and we will report the R^2 .

The critical values for the first-stage F -statistics are displayed in Table 5.2 in Stock & Yogo (2005). Thus, the values should not be below $F = 5.53$ (H2, H3, and H4).

AP8 Exploratory analysis (optional)

Describe any exploratory analyses to be conducted with your data. Include here any planned analyses that are not confirmatory in the sense of being a direct test of one of the specified hypotheses.

We will test whether there are differences between the neutral and the servant leadership speech regarding charismatic leadership perceptions using OLS.

AP9 Other information (optional)

Other information optional

(NOTE: If needed, multiple lines with other information can be included)

O1 Other information (optional)

If there is any additional information that you feel needs to be included in your preregistration, please enter it here. Literature cited, disclosures of any related work such as replications or work that uses the same data, or other context that will be helpful for future readers would be appropriate here.

References

R1 References

Enter your references below. Use a consistent format (e.g., <https://apastyle.apa.org/style-grammar-guidelines/references/examples>)

- Alvesson, M. (2020). Upbeat leadership: A recipe for – or against – “successful” leadership theory. *The Leadership Quarterly*, 31(6). <https://doi.org/10.1016/j.leaqua.2020.101439>
- Alvesson, M., & Einola, K. (2019). Warning for excessive positivity: Authentic leadership and other traps in leadership studies. *The Leadership Quarterly*, 30(4), 383–395. <https://doi.org/10.1016/j.leaqua.2019.04.001>
- Anderson, M. H., & Sun, P. Y. T. (2017). Reviewing leadership styles: Overlaps and the need for a new ‘full-range’ theory. *International Journal of Management Reviews*, 19(1), 76–96. <https://doi.org/10.1111/ijmr.12082>
- Antonakis, J. (2017). On doing better science: From thrill of discovery to policy implications. *The Leadership Quarterly*, 28(1), 5–21. <https://doi.org/10.1016/j.leaqua.2017.01.006>
- Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly*, 21(6), 1086–1120. <https://doi.org/10.1016/j.leaqua.2010.10.010>
- Antonakis, J., d’Adda Giovanna, Weber, R., & Zehnder, C. (2015). *Just words? Just speeches? On the economic value of charismatic leadership*. NBER Rep. 4. http://www.hec.unil.ch/jantonakis/adwz_charisma.pdf
- Antonakis, J., d’Adda, G., Weber, R., & Zehnder, C. (2021). Just words? Just speeches? On the economic value of charismatic leadership. *Management Science*.
- Bande, B., Fernández-Ferrín, P., Varela-Neira, C., & Otero-Neira, C. (2016). Exploring the relationship among servant leadership, intrinsic motivation and performance in an industrial sales setting. *Journal of Business and Industrial Marketing*, 31(2), 219–231. <https://doi.org/10.1108/JBIM-03-2014-0046>
- Bandura, A. (1977). *Social learning theory*. Prentice-Hall.
- Banks, G. C., Gooty, J., Ross, R. L., Williams, C. E., & Harrington, N. T. (2018). Construct redundancy in leader behaviors: A review and agenda for the future. *The Leadership Quarterly*, 29(1), 236–251. <https://doi.org/10.1016/j.leaqua.2017.12.005>
- Banks, G. C., McCauley, K. D., Gardner, W. L., & Guler, C. E. (2016). A meta-analytic review of authentic and transformational leadership: A test for redundancy. *The Leadership Quarterly*, 27(4), 634–652. <https://doi.org/10.1016/j.leaqua.2016.02.006>
- Bass, B. M. (1999). Two decades of research and development in transformational leadership. *European Journal of Work and Organizational Psychology*, 8(1), 9–32. <https://doi.org/10.1080/135943299398410>
- Bernardin, H. J., Cooke, D. K., & Villanova, P. (2000). Conscientiousness and agreeableness as predictors of rating leniency. *Journal of Applied Psychology*, 85(2), 232–236. <https://doi.org/10.1037/0021-9010.85.2.232>
- Blau, P. M. (1964). *Exchange and power in social life*. Wiley.
- Bono, J. E., Hooper, A. C., & Yoon, D. J. (2012). Impact of rater personality on transformational and transactional leadership ratings. *The Leadership Quarterly*, 23(1), 132–145. <https://doi.org/10.1016/j.leaqua.2011.11.011>
- Butler, D., & Geis, F. L. (1990). Nonverbal affect responses to male and female leaders: Implications for leadership evaluations. *Journal of Personality and Social*

- Psychology*, 58(1), 48–59. <https://doi.org/10.1037/0022-3514.58.1.48>
- Chen, Z., Zhu, J., & Zhou, M. (2015). How does a servant leader fuel the service fire? A multilevel model of servant leadership, individual self identity, group competition climate, and customer service performance. *The Journal of Applied Psychology*, 100(2), 511–521. <https://doi.org/10.1037/a0038036>
- Choudhary, A. I., Akhtar, S. A., & Zaheer, A. (2013). Impact of transformational and servant leadership on organizational performance: A comparative analysis. *Journal of Business Ethics*, 116(2), 433–440. <https://doi.org/10.1007/s10551-012-1470-8>
- Demetriou, A., Kyriakides, L., & Avraamidou, C. (2003). The missing link in the relations between intelligence and personality. *Journal of Research in Personality*, 37(6), 547–581. [https://doi.org/10.1016/S0092-6566\(03\)00048-5](https://doi.org/10.1016/S0092-6566(03)00048-5)
- Eva, N., Robin, M., Sendjaya, S., Van Dierendonck, D., & Liden, R. C. (2019). Servant leadership: A systematic review and call for future research. *The Leadership Quarterly*, 30(1), 111–132. <https://doi.org/10.1016/j.leaqua.2018.07.004>
- Felfe, J., & Heinitz, K. (2010). The impact of consensus and agreement of leadership perceptions on commitment, organizational citizenship behaviour, and customer satisfaction. *European Journal of Work and Organizational Psychology*, 19(3), 279–303. <https://doi.org/10.1080/13594320802708070>
- Furnham, A., Moutafi, J., & Chamorro-Premuzic, T. (2005). Personality and intelligence: Gender, the big five, self-estimated and psychometric intelligence. *International Journal of Selection and Assessment*, 13(1), 11–24. <https://doi.org/10.1111/j.0965-075X.2005.00296.x>
- Gardner, W. L., Lowe, K. B., Meuser, J. D., Noghani, F., Gullifor, D. P., & Cogliser, C. C. (2020). The leadership trilogy: A review of the third decade of the leadership quarterly. *The Leadership Quarterly*, 31(1), 101379. <https://doi.org/10.1016/j.leaqua.2019.101379>
- Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. de Fruyt, & F. Ostendorf (Eds.), *Personality Psychology in Europe* (pp. 7–28). Tilburg University Press.
- Gotsis, G., & Grimani, K. (2016). The role of servant leadership in fostering inclusive organizations. *Journal of Management Development*, 35(8), 985–1010. <https://doi.org/10.1108/JMD-07-2015-0095>
- Greenleaf, R. K. (1977). *Servant-leadership: A journey into the nature of legitimate power and greatness*. Paulist Press.
- Güntner, A. V., Klonek, F., Lehmann-Willenbrock, N., & Kauffeld, S. (2020). Follower behavior renders leader behavior endogenous: The simultaneity problem, estimation challenges, and solutions. *The Leadership Quarterly*, 31(6). <https://doi.org/10.1016/j.leaqua.2020.101441>
- Hansbrough, T. K., Lord, R. G., & Schyns, B. (2015). Reconsidering the accuracy of follower leadership ratings. *The Leadership Quarterly*, 26(2), 220–237. <https://doi.org/10.1016/j.leaqua.2014.11.006>
- Hoch, J. E., Bommer, W. H., Dulebohn, J. H., & Wu, D. (2018). Do ethical, authentic, and servant leadership explain variance above and beyond transformational leadership? A meta-analysis. *Journal of Management*, 44(2), 501–529. <https://doi.org/10.1177/0149206316665461>
- Khademi, M., Schmid Mast, M., Zehnder, C., & Saint Priest, O. de (2021). The problem of demand effects in power studies: Moving beyond power priming. *The Leadership Quarterly*, 32(4), 101496. <https://doi.org/10.1016/j.leaqua.2021.101496>
- Lee, A., Lyubovnikova, J., Tian, A. W., & Knight, C. (2020). Servant leadership: A meta-analytic examination of incremental contribution, moderation, and mediation. *Journal of Occupational and Organizational Psychology*, 93(1), 1–44. <https://doi.org/10.1111/joop.12265>
- Lemoine, G. J., & Blum, T. C. (2021). Servant leadership, leader gender, and team gender

- role: Testing a female advantage in a cascading model of performance. *Personnel Psychology*, 74(1), 3–28. <https://doi.org/10.1111/peps.12379>
- Lemoine, G. J., Hartnell, C. A., & Leroy, H. (2019). Taking stock of moral approaches to leadership: An integrative review of ethical, authentic, and servant leadership. *Academy of Management Annals*, 13(1), 148–187. <https://doi.org/10.5465/annals.2016.0121>
- Lehrke, A. S., & Sowden, K. (2017). Servant leadership and gender. In C. J. Davis (Ed.), *Servant leadership and followership: examining the impact on workplace behavior* (pp. 25–50). Springer International Publishing. <https://doi.org/10.1007/978-3-319-59366-1>
- Liden, R. C., Wayne, S. J., Zhao, H., & Henderson, D. (2008). Servant leadership: Development of a multidimensional measure and multi-level assessment. *The Leadership Quarterly*, 19(2), 161–177. <https://doi.org/10.1016/j.leaqua.2008.01.006>
- Lonati, S., Quiroga, B. F., Zehnder, C., & Antonakis, J. (2018). On doing relevant and rigorous experiments: Review and recommendations. *Journal of Operations Management*, 64(1), 19–40. <https://doi.org/10.1016/j.jom.2018.10.003>
- Luthans, F., & Avolio, B. J. (2003). Authentic leadership development. In K. S. Cameron, J. E. Dutton, & R. E. Quinn (Eds.), *Positive organizational scholarship* (pp. 241–258). Barrett-Koehler.
- Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods*, 17(3), 437–455. <https://doi.org/10.1037/a0028085>
- Meslec, N., Curseu, P. L., Fodor, O. C., & Kenda, R. (2020). Effects of charismatic leadership and rewards on individual performance. *The Leadership Quarterly*, 31(6). <https://doi.org/10.1016/j.leaqua.2020.101423>
- O'Connor, M. C., & Paunonen, S. V. (2007). Big five personality predictors of post-secondary academic performance. *Personality and Individual Differences*, 43(5), 971–990. <https://doi.org/10.1016/j.paid.2007.03.017>
- Orbell, S., & Sheeran, P. (1998). 'Inclined abstainers': A problem for predicting health-related behaviour. *The British Journal of Social Psychology*, 37(2), 151–165. <https://doi.org/10.1111/j.2044-8309.1998.tb01162.x>
- Parris, D. L., & Peachey, J. W. (2013). A systematic literature review of servant leadership theory in organizational contexts. *Journal of Business Ethics*, 113, 377–393. <https://doi.org/10.1007/s10551-012-1322-6>
- Pircher Verdorfer, A., & Peus, C. (2014). The measurement of servant leadership: validation of a german version of the servant leadership survey (sls). *Zeitschrift Für Arbeits- Und Organisationspsychologie*, 58, 1–16. <https://doi.org/10.1026/0932-4089/a000133>
- Podsakoff, P. M., & Podsakoff, N. P. (2019). Experimental designs in management and leadership research: Strengths, limitations, and recommendations for improving publishability. *The Leadership Quarterly*, 30(1), 11–33. <https://doi.org/10.1016/j.leaqua.2018.11.002>
- Sajons, G. B. (2020). Estimating the causal effect of measured endogenous variables: A tutorial on experimentally randomized instrumental variables. *The Leadership Quarterly*, 31(5). <https://doi.org/10.1016/j.leaqua.2019.101348>
- Semadeni, M., Withers, M. C., & Trevis Certo, S. (2014). The perils of endogeneity and instrumental variables in strategy research: Understanding through simulations. *Strategic Management Journal*, 35(7), 1070–1079. <https://doi.org/10.1002/smj.2136>
- Schyns, B., & Sanders, K. (2007). In the eyes of the beholder: Personality and the perception of leadership. *Journal of Applied Social Psychology*, 37(10), 2345–2363. <https://doi.org/10.1111/j.1559-1816.2007.00261.x>
- Sousa, M., & Van Dierendonck, D. (2017). Servant leaders as underestimators: Theoretical and practical implications. *Leadership & Organization Development Journal*, 38(2), 270–283. <https://doi.org/10.1108/LODJ-10-2015-0236>
- Stock, J. H., & Yogo, M. (2005). Testing for weak instruments in linear iv regression. In D.

- W. K. Andrews, J. H. Stock, & T. J. Rothenberg (Eds.), *Identification and inference for econometric models: essays in honor of Thomas Rothenberg* (pp. 80–108). Cambridge University Press. <https://doi.org/10.1017/CBO9780511614491.006>
- Sutin, A. R., Stephan, Y., Luchetti, M., Strickhouser, J. E., Aschwanden, D., & Terracciano, A. (2021). The association between five factor model personality traits and verbal and numeric reasoning. *Neuropsychology, Development, and Cognition*, 1–21. <https://doi.org/10.1080/13825585.2021.1872481>
- Tajfel, H. (1978). Differentiation between social groups: Studies in the social psychology of intergroup relations. In H. Tajfel (Ed.), *Differentiation between social groups: Studies in the social psychology of intergroup relations*. Academic Press.
- Van Dierendonck, D., & Nuijten, I. (2011). The servant leadership survey: Development and validation of a multidimensional measure. *Journal of Business and Psychology*, 26(3), 249–267. <https://doi.org/10.1007/s10869-010-9194-1>
- Van Dierendonck, D., Stam, D., Boersma, P., de Windt, N., & Alkema, J. (2014). Same difference? Exploring the differential mechanisms linking servant leadership and transformational leadership to follower outcomes. *The Leadership Quarterly*, 25(3), 544–562. <https://doi.org/10.1016/j.leaqua.2013.11.014>
- Wang, G., van Iddekinge, C. H., Zhang, L., & Bishoff, J. (2019). Meta-analytic and primary investigations of the role of followers in ratings of leadership behavior in organizations. *Journal of Applied Psychology*, 104(1), 70–106. <https://doi.org/10.1037/apl0000345>
- Zizzo, D. J. (2010). Experimenter demand effects in economic experiments. *Experimental Economics*, 13(1), 75–98. <https://doi.org/10.1007/s10683-009-9230-z>

This document was created using the **Psychological Research Preregistration-Quantitative (aka PRP-QUANT) Template**, version 2 (available at <https://www.psycharchives.org/>).

The template was developed by a task force composed of members of the American Psychological Association (APA), the British Psychological Society (BPS), the German Psychological Society (DGPs), the Center for Open Science (COS), and the Leibniz Institute for Psychology (ZPID). This work is licensed under the [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/) license. Thus, you are free to share and adapt the content, given that you attribute the source and indicate if changes were made.

The implementation as Google Doc was done by ZPID. Find out more about ZPID and our preregistration service **PreReg** by visiting <https://leibniz-psychology.org/> and <http://prereg-psych.org/>, respectively.

To receive a timestamp and a DOI (digital object identifier), submit your preregistration protocol to **PsychArchives** via <https://pasa.psycharchives.org/>, preferably as PDF.