
2D or 3D social evaluation? A pre-registered, cross-cultural, multi-site study

Keywords: Person perception, social evaluation, warmth, competence, morality.

Last updated: December 4th 2024

Abstract

Recent adversarial collaborations on person perception, and person and group stereotypes suggest two main dimensions of social evaluation (Horizontal dimension: Warmth/Communion: and Vertical dimension: Competence/Agency). However, research focused on moral character suggests a tri-dimensional model where morality is a third and more central dimension of person perception that can be distinguished from the horizontal dimension. This pre-registered multi-site study aims to directly contrast these hypotheses across a number of social targets, including the self, another agent, an in-group, an out-group, and agents from in-groups and out-groups. We will do this by comparing goodness-of-fit of a number of confirmatory factor models instantiating either bi-dimensional or tri-dimensional models overall and across different social targets. Exploratory analysis will deal with possible culture-level variables at play in favouring either model.

Research Transparency Statement

Conflicts of interest: None of the authors have any conflict of interest to declare.

Funding: At the time of stage 1 submission, no funding has been granted to this project. We plan to submit to funding opportunities and will declare them if needed.

Artificial intelligence: No artificial intelligence has been used in any stage of this project.

Ethics: Project has been approved by Universidad de los Andes (Bogotá – Colombia) IRB.

Computational reproducibility: Authors are aiming for a computational reproducibility badge. R data analysis scripts are freely available on the project's OSF page, and any change to them between stages 1 and 2 will be clearly marked.

Data: As this is a stage 1 registered report, no data have been collected. We will make all raw and edited data freely available on the project's OSF page. We are applying for open data, open materials, and preregistration badges.

Materials: English materials are freely available on the project's OSF page and all translated materials will be freely available there too. We are applying for open data, open materials, and preregistration badges.

Analysis scripts: R data analysis scripts are freely available in the project's OSF page. Any change to them between stages 1 and 2 will be clearly marked. We are applying for open data, open materials, and preregistration badges.

Pre Published Manuscript DO NOT REPRODUCE

2d or 3d social evaluation? A pre-registered, multi-site, multi-cultural study

Social evaluation is the process through which people generate impressions and expectations of other agents or groups and guide their actions in consequence. Classic studies in social psychology have aimed to identify the fundamental dimensions of social evaluation (Abele & Wojciszke, 2007; Brambilla, Sacchi, Rusconi, & Goodwin, 2021; Cuddy, Fiske, & Glick, 2008; Ellemers & van den Bos, 2012; Yzerbyt, 2018). Literature suggests several models differing on *I*) the target of evaluation (e.g., a person or group vs. the self), *II*) the number of dimensions (e.g. two or three?), and *III*) the dimensions themselves (i.e., Warmth and Competence vs. Agency and Communion) (Abele et al, 2020; Koch et al, 2021).

First, the Stereotype Content Model (SCM; Cuddy, Fiske, & Glick, 2008; Fiske, Cuddy, Glick, & Xu, 2018) and the Dimensional Compensation Model (DCM; Yzerbyt, 2018; Yzerbyt, Kervyn, & Judd, 2008) posit two dimensions of social evaluation: Warmth and Competence. Warmth reflects how much the target's intentions are perceived to be aligned with the self, while Competence deals with whether the target is perceived to be able to carry out their intentions, whatever those are. The Dual Perspective Model (DPM; Abele & Wojciszke, 2007, 2014) follows a similar bi-dimensional structure of Agency and Communion, instead of Warmth and Competence. Agency includes qualities relevant to attaining personal goals, whatever those may be and Communion are qualities relevant to building social relationships. Further, this model distinguishes two facets related to each dimension with Agency being composed of Assertiveness and Ability and Communion of Morality and Friendliness (Abele et al., 2016, Fiske et al, 2021). For these models, 'moral' traits such as honesty are included in the Warmth/Communion dimension, along with "social" traits such as being funny or cooperative.

This bi-dimensional picture, however, is disputed. The Behavioural Regulation Model (BRM; Ellemers & van den Bos, 2012) and Moral Character research (MC; Brambilla, Sacchi, Rusconi, & Goodwin, 2021; Goodwin, Piazza, & Rozin, 2014; Pizarro & Tannenbaum, 2011; Uhlmann, Pizarro, & Diermeier, 2015) suggest that Morality constitutes a third fundamental dimension of social evaluation distinct from but equally important as Warmth/Communion and Competence/Agency. Here, Morality reflects whether the judged agent appears to be a willing and consistent cooperator, whereas Warmth ratings reflect whether social dealings with them are pleasurable independently of whether they are actually cooperating. This tri-dimensional model is echoed by the Agency, Beliefs and Communion Model (ABC; Koch, Imhoff, Dotsch, Unkelbach, & Alves, 2016). Much like prior models, ABC suggests that the Morality dimension is independent and at least as important as the two others. For the ABC model, Beliefs dimension reflects whether the agent holds more socially liberal or conservative values and could be analogous to the Moral dimension inasmuch as two people holding similar beliefs (e.g., two highly conservative people) would presumably deem each other to be morally right whereas people having very dissimilar beliefs (e.g., a very liberal person and a very conservative person) would deem each other to be immoral.

A recent adversarial collaboration (Abele, Ellemers, Fiske, Koch, & Yzerbyt, 2020; Koch, Yzerbyt, Abele, Ellemers, & Fiske, 2021) tried to unify these models. Adversarial collaboration takes a theoretical approach to smooth over apparent dissimilarities across models around the number, content and relationship between fundamental dimensions, and whether they change according to social target (e.g., the self vs. another person). It concludes that there are two fundamental dimensions: a vertical dimension integrating Agency and Competence and a horizontal dimension integrating Communion, Warmth, Moral, Friendliness, and Belief. The

vertical dimension is divided into Ability and Assertiveness facets and the horizontal dimension into Friendliness and Morality (Abele et al., 2021, p. 297) (see Figs 1 – 2).

Adversarial Collaboration takes at face value the unification of warm (e.g., funny) and moral traits (e.g., honest) into a single dimension with limited direct empirical evidence and contrary to prior evidence suggesting Morality as a third, equally important dimension (Brambilla et al., 2021; Critcher, Helzer, & Tannenbaum, 2020; Ellemers & van den Bos, 2012; Goodwin et al., 2014; Leach, Ellemers, & Barreto, 2007; Pizarro & Tannenbaum, 2011; Uhlmann et al., 2015). Literature suggests that Morality is more central in people's overall social perceptions, above and beyond warmth and competence attributions (Brambilla & Leach, 2014; Brambilla, Rusconi, Sacchi, & Cherubini, 2011; Brambilla et al., 2021). Similar results apply to other aspects of social cognition such as liking or respecting (Hartley et al., 2016), true self-attribution (Newman, Bloom, & Knobe, 2014; Strohminger, Knobe, & Newman, 2017), sense of self (Stanley et al., 2019; Stanley, Henne, Brigard, & Stanley, 2018) and social desirability (Landy, Piazza, & Goodwin, 2016). Overall, this line of research supports a tri-factorial model in which a Morality dimension is at least equally, if not more central, than the horizontal and vertical dimensions (Brambilla et al., 2021) (see fig 3). In an effort to bridge together bi- and tri-dimensional models, we propose a tri-dimensional “complex” model (fig 4). In this tri-dimensional complex model, we apply the concept of facets proposed by adversarial collaboration by dividing the vertical dimension into Ability and Assertiveness, but keeping Friendliness and Morality apart as full-fledged, independent dimensions, rather than as facets of the horizontal dimension.

While we celebrate and encourage the willingness of adversarial collaboration authors to rigorously develop and unify apparently contradictory theories (Abele et al., 2020; Koch et al.,

2021), we believe their conclusions require additional direct empirical support. So far, the unification of these theories results from this contribution stem from theoretical reflection rather than a direct, empirical confrontation of proposed models. Indeed, bi-dimensional models have received empirical support in a variety of cultural contexts (Abele et al, 2016; Cuddy et al., 2008; Durante et al., 2013; Fiske et al., 2007, 2018). Nevertheless, these studies have only considered and analysed the goodness-of-fit of bi-dimensional models to data, seldom comparing model fit to competing models featuring other factorial structures (Sayans-Jiménez et al, 2017). Therefore, they cannot offer evidence on whether a bi-dimensional model exhibits a better fit *compared to* an alternative, theoretically relevant model.

A remarkably limited number of investigations have employed Confirmatory Factor Analysis (CFA) or any similar hypothesis-confronting strategy to directly compare candidate models. Abele et al. (2016) found cross-cultural evidence of a good fit for a two-factor solution compared to a four-factor solution. This bi-dimensional solution showed somewhat acceptable fit statistics but was not compared to competing models both in cross-cultural testing (Kosakowska-Berezecka et al., 2023) and using a meta-analytic approach (Friebs et al, 2022). Similarly, Carrier et al. (2014) examined whether Agency and Competence dimensions are the same by comparing a two-factor solution uniting Agency and Competence/Warmth in a second factor, against a three-factor solution where all dimensions fit independently. Surprisingly, the latter had the best fit, suggesting that Agency and Competence dimensions might be distinct factors. In the largest study on the matter to date, authors validated a two-item measure of proposed facets within horizontal and vertical dimensions compared to a 5-item per facet measure (Koch et al., n.d.). They observed that a two-item per facet measure offers better data fit than 5-item measures for both bi-dimensional “simple” models (i.e., not distinguishing facets

within dimensions) and bi-dimensional “complex” models (i.e., distinguishing facets within dimensions). Yet, the authors did not fit a three-dimensional model distinguishing Morality and Friendliness as distinct but equally important dimensions. In this sense, while psychometrically valid, this study does not offer evidence pertaining to possible three-dimension models distinguishing Morality and Friendliness as equally important dimensions.

A relatively small number of studies have directly compared bi- and tri-dimensional models. Leach et al. (2007) compared the fit of the three-dimensional solution (BRM: Competence, Warmth and Morality) against a competing bi-dimension solution unifying Morality and Warmth. The tri-dimensional solution was by far the best fitting model. Similar results were found dealing with group stereotypes (López-Rodríguez, Cuadrado, & Navas, 2013). However, results on region and country stereotypes fail to find conclusive evidence when directly comparing bi-dimensional and tri-dimensional models (Barbosa et al., n.d.).

In summary, even though there is ample evidence in favour of bi-dimensional models, direct contrast of bi-dimensional models to plausible alternatives has seldom been implemented and in the few cases it has, no conclusive evidence for or against either model was obtained (Barbosa et al., n.d.; Carrier, Louvet, Chauvin, & Rohmer, 2014; López-Rodríguez et al., 2013; Wayne Leach, Ellemers, & Barreto, 2007). Moreover, when CFA comparison of different models is used, direct confrontation of competing theories can feature ad hoc models that either reflect different aspects of the same theory (e.g., Abele et al., 2016) or a hybrid model that does not reflect any established theory (e.g., Carrier et al., 2014). These analytical strategies make it impossible to make empirically informed decisions on which (if any) model ought to be preferred (Crede & Harms, 2019).

Thus, our multi-site, pre-registered report aims to cross-culturally determine which between the bi-factorial “simple” (fig 1), bi-factorial “complex” (fig 2), tri-factorial “simple” or “complex” models (figs 3 - 4) better describes social evaluation across targets. All theoretically relevant models will be compared to two theoretically irrelevant models, a uni-dimensional control model suggesting no underlying factors of social evaluation but rather a single continuum (fig 5) and a 4D model (fig 6) claiming that all four facets of social evaluation are independent factors rather than two sub-factors (i.e., Ability and Assertiveness, and Sociability and Morality for Vertical and Horizontal dimensions respectively).

All CFA model comparisons will be done using fit (Chi², RMSEA, CFI, TLI, and SRMS) and generalizability indexes (AIC and BIC) computed and analysed for all models to determine which of them better fits social evaluation across different social targets. Overall, we aim to offer culturally wide evidence in favour of one model for solid, empirically-backed further research on social evaluation. While we do not believe our study can claim an exhaustive list of all relevant dimensions of person perception (Hester & Hehman, 2020; Naumann, Vazire, Rentfrow, & Gosling, 2009; Nicolas, Bai, & Fiske, 2022), we do claim that our results could offer general insights into the most central dimensions and whether they change according to social target.

Fig 1. Bi-factorial "simple" model.

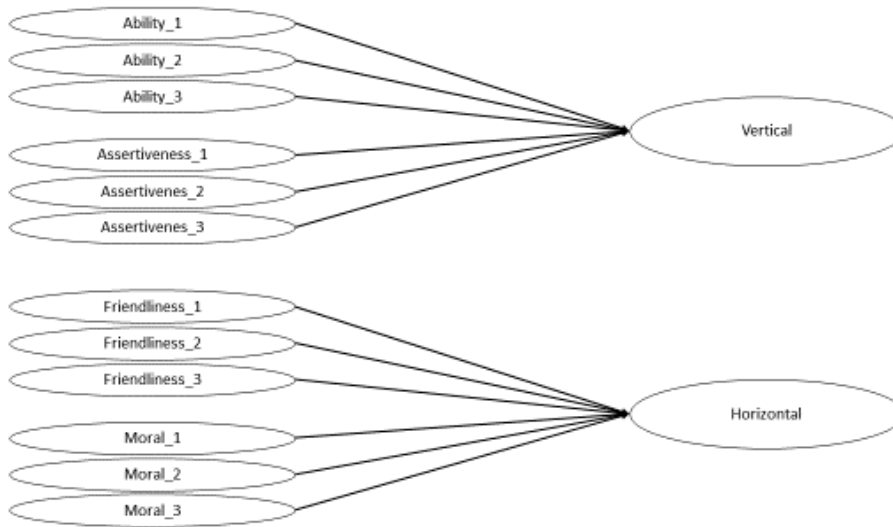


Fig 2. Bi-factorial "complex" model.

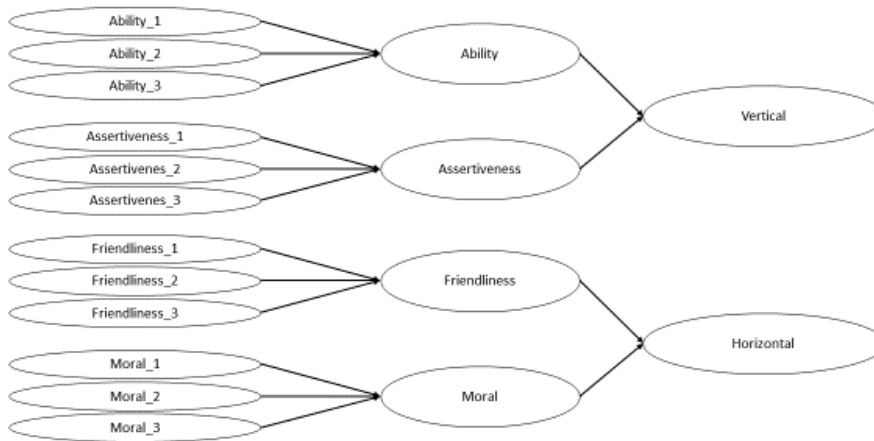


Fig 3. Tri-factorial "simple" model.

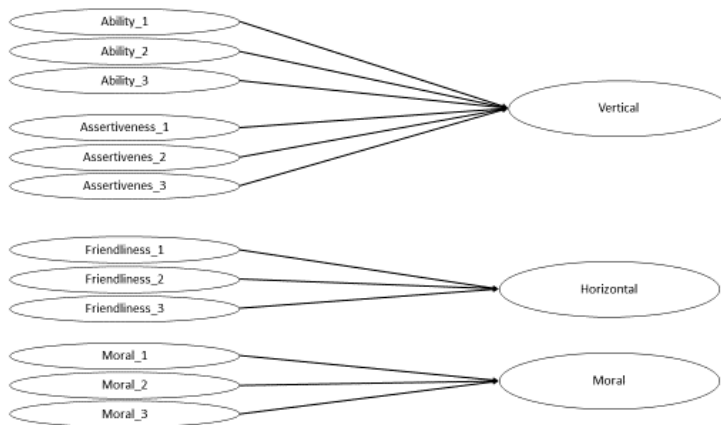


Fig 4. Tri-factorial "complex" model.

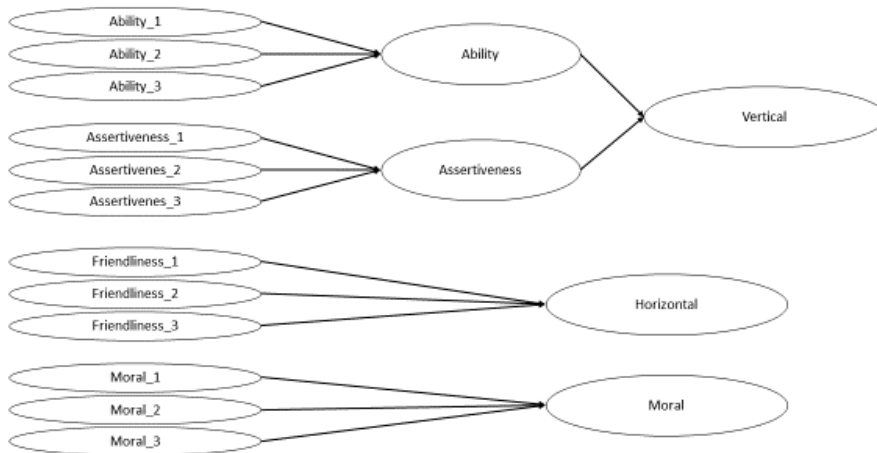


Fig 5. Uni-factorial "control" model.

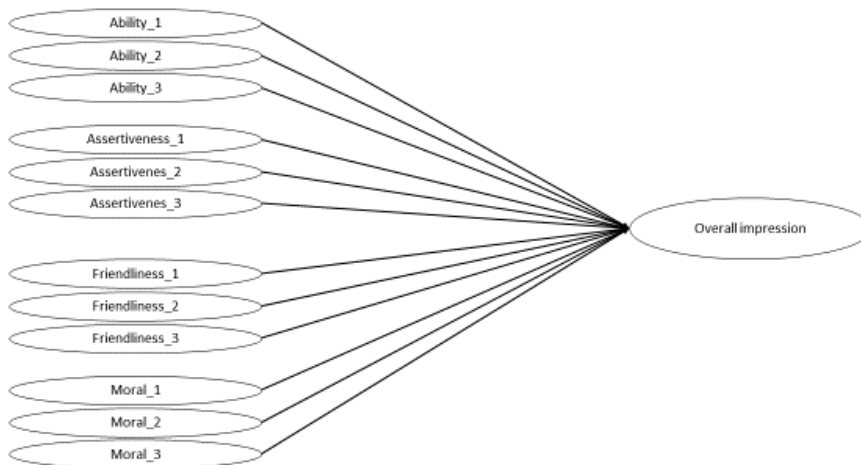
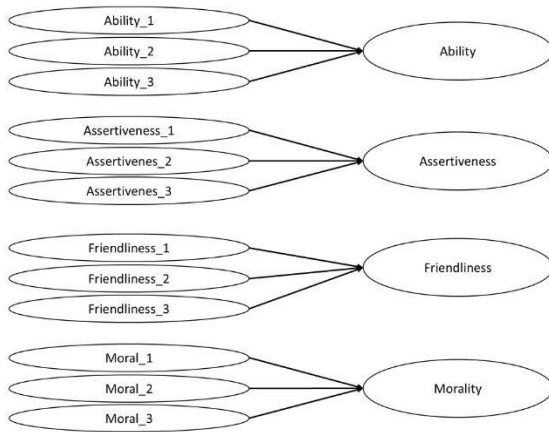


Fig 6. 4D control model.



Social Target as a Moderator of Model Fit

One of the main contributions of the adversarial collaboration was to demarcate the aim and scope of each model, highlighting that each one focused on evaluating different social targets (Abele et al., 2020; Koch et al., 2021). For example, while DPM concentrates on the evaluation of self vs. others, BRM and DCM focus on in-group and out-group evaluation, respectively, and SCM and ABC are centred on evaluations of groups, not individual people, regardless of their social group. This means that each theory has been built to understand how people assess different types of targets, and hence their assertions might be limited to them. Adversarial collaboration attempted to resolve this by proposing the type of target being evaluated as a moderator of the used dimensional structure and that people might prioritise horizontal, vertical, or even additional dimensions depending on this. Thus, each model could explain the social evaluation of specific targets more effectively and therefore have varying fit according to the considered social target. For instance, both Abele et al. (2021) and Koch et al. (2021) believe the Morality dimension separates from the Warmth dimension only when people evaluate other in-group members, but not when they evaluate out-group members. On the contrary, Moral

Character (MC) predicts that the Morality dimension should be clearly distinguished from other dimensions for all social targets (Brambilla et al, 2021; Goodwin et al, 2014). Hence, we will run pre-registered analyses fitting all presented models distinguishing between different social targets: the self, another person, an ingroup (e.g., latinx people if the self *is* latinx), an outgroup (e.g., latinx people if the self is *not* latinx), another person belonging to an ingroup (e.g., Juana, a latina woman, if the self *is* latinx) and another person belonging to an outgroup (e.g., Juana, a latina woman, if the self is *not* latinx). We have no specific prediction on this comparison but will run identical analyses for each social target and report whether the preferred model changes across them.

Cross-cultural Variation

Psychological research heavily relies on WEIRD samples (Henrich et al., 2010), which may limit the generalizability of findings. Here, these variations may be explained by cultural influence on cognitive and social processes, leading to differences in social evaluation (Henrich, 2020; Nisbett et al., 2001; Norenzayan et al., 2002); and also by cultural variation in values influencing the traits reflecting each dimension across cultures (Butchel et al., 2018). These cultural differences, coupled with discussions on the replicability of results on non-WEIRD samples (Frimer, 2020; Kivikangas et al., 2021), underscore the limitations of generalising findings across diverse cultural contexts (Davis et al., 2016; Iurino & Saucier, 2020). In that sense, cross-cultural empirical comparisons become essential to overcome these limitations.

We will conduct Multi Level Confirmatory Factor Analysis (MLCFA) to empirically compare the dimensional structure of folk person perception across different cultures, exploring the inclusion of country-level variables as predictors of preferred models. These variables may include the Gini Index, Gender Inequality Index, Subjective Well-Being, Multidimensional

Poverty Index, GDP per capita, collectivism-individualism, religiosity, Human Development Index, percentage of Urban Population, and Life Expectancy at birth. While we do not have specific directional hypotheses for these variables, we believe they can offer insights into how societal-level variables may influence social perception and will run and report exploratory analyses on them.

Method

Site Recruitment

We have informally contacted potential allied labs worldwide through mailing lists ascribed to psychological science organisations (e.g., Judgment and Decision Making, Center for the Science of Moral Understanding, APA, Psychological Science Accelerator, ABRIR – PSY, Open Science Foundation, Society for the Improvement of Psychological Science...) as well as relevant platforms on social media (e.g., Psychological Research Methods and Moral Psychology groups on Facebook, suitable hashtags on Twitter...) and co-authors' own personal networks¹. This strategy will be continued throughout the review process to reach a large and wide number of allied authors across the globe. All interested labs or researchers with a proven record of research in psychological science and a university affiliation will be considered for participation and will receive authorship in the final manuscript provided they can *i*) get IRB approval for data collection in their country, *ii*) adapt all materials to their particular sociocultural context, *iii*) collect data from at least 250 valid participants before the end of the data collection phase (see below in online supplementary materials for inclusion/ exclusion criteria, a priori power analysis

¹ As of July 3rd 2024 there are 90 people interested across 44 different countries. They have not been shown the present manuscript yet but we expect them and hopefully more authors to include the final list of authors. Requirements to be an author in the final manuscript are presented below.

and expected data analysis), and *iv*) send all required materials and raw data to the project's OSF page, including translation and back-translations and data quality validations.

We are aware that similar open call projects show an overrepresentation of universities in the US and Western Europe given that these countries tend to have the resources and infrastructure to rapidly collect data. Therefore, we also aim to collect data along two sources of variation: Global North and Global South, and at least two labs per continent. We will do this with targeted calls to labs in these areas. While this is not a guarantee of participation, we will strive to have an overall sample that reflects adequate cross-cultural variation.

As soon as IPA is obtained, the data collection phase with participating labs will start. In this phase, depending on the lab's local university policy, labs will get on-site IRB approval or will use main IRB approval by [REDACTED FOR PEER REVIEW]. In any case each lab will be expected to adapt all materials, and recruit a minimum of 250 participants. Also, all allied labs will provide detailed logs of their data collection procedures, such as lab notes, time, place, and the context in which data collection was carried out as well as keep contact with the corresponding author's team. Data collection will be done through a centralized Qualtrics online survey adapted for each country/language. We expect each lab to take approximately 6 months for data collection, including obtaining IRB approval. Yet the data collection phase will last for a full year as of IPA to account for delays in IRB approval, local logistics, data collection, and validation. After the data collection phase, all raw data files will be compiled and analysed centrally as described below.

Expected Sample

Main Analysis

We will fit five CFA models reflecting bi-factorial “simple” and “complex” models, tri-factorial “simple” and “complex” models, as well as a theoretically irrelevant uni-factorial model (see figs 1 - 5) (see lavaan code for each model below). We ran an a priori power analysis for each of these models assuming $\Delta RMSEA = 0.07$, $\alpha = 0.05$, and $\beta = 0.99$ using the semPower package in R (Jobst, Bader, & Moshagen, 2021). Results suggested a minimum sample of 229 to accurately observe any of the proposed models. We estimate an average sample attrition of 10% due to fitting exclusion criteria (see below). Therefore, each partner lab will be required to collect a minimum of $n = 250$ *after inclusion/ exclusion criteria* to allow the accurate fitting of all considered CFA models for each lab. According to their own possibilities, each lab will collect samples for either convenience or representative samples. Full demographic characterization of participants (i.e., age, gender, educational level, religiosity, and self-reported political preferences) will be asked after data collection².

Materials and Procedure

Cultural Adaptation Procedure

The participating laboratories will translate the survey items into the language of the participant pool, following the translation process detailed below:

1. Translation: The original document is translated from the source to the target language by translators resulting in document version A.
2. Back-translation: Version A is translated back from the target to the source language by B translators independently, resulting in version B.

² Demographics questions may be added or removed according to each countries' specific circumstances and in accordance with specific IRB requirements. Demographic data will be used only as sample characterization; no pre-registered hypothesis deals with any demographic variable.

3. Discussion: Version A and B are discussed among translators and the language coordinator, discrepancies in version A and B are detected and solutions are discussed. Version C is created.
4. External readings: Version A is read and evaluated by external reviewers to ensure comprehension and validity.
5. Cultural adjustments: minor details such as names, culturally sensitive topics, and idiosyncratic vocabulary can be edited.

Data Collection Procedure

Data collection will be run through a single online Qualtrics survey for each country. Participants will complete an informed consent and then start main data collection followed by a short demographic questionnaire. Allied labs can include other measures in this data collection provided they are implemented *after* the main data collection is done.

In the main data collection phase, participants will be asked to think about and briefly describe a social target. We will randomly assign each participant to three social targets out of six (Target variable, between-subjects independent variable with six levels: the self, another person, an in-group, an outgroup, another person from an in-group, and another person from an outgroup). To ensure variability in ratings, participants will also be randomly assigned to a single experimental condition dealing with whether the target is liked, disliked, or neutral (Condition variable, between-subjects independent variable with three levels: like, dislike, and neutral). Notably, all three randomly assigned targets will be framed in the same Condition variable level. For instance, a participant randomly assigned to judge another person, the in-group, and the outgroup in the dislike condition will be asked to think of a person they dislike (combination of Target: Person and Condition: Dislike), a group they belong to but are not particularly fond of

and would rather not belong to (combination of Target: in-group and Condition: Dislike), and an outgroup they dislike (combination of Target: outgroup and Condition: Dislike)³.

In all conditions, participants will be asked to think of a social target, briefly describe them, and then deal out personality trait attributions, willingness to engage in different types of social relationships and emotions associated with the target. Description of the social target, whether a single agent or a group, allows us to ensure that participants are thinking about a specific, real target as well as to ensure attention and data quality. Open-ended descriptions will be dealt with in the dominant language of the country data is being collected in (e.g., Spanish in all Latin American countries except Brazil, French in France and other francophone countries, English in the US, UK, Australia...) and will be verified by each lab to ensure that participants followed the instructions. These descriptions will be used as exclusion criteria whereby participants failing to follow instructions for this manipulation will be excluded. Each lab will be expected to review all open-ended descriptions and exclude short or overly generic descriptions as well as those who failed to comply with instructions (e.g., giving positive descriptions of a target in the Dislike condition or describing a single person in the in-group condition).

After thinking about and describing each target, participants will be asked to rate the target's friendliness, morality, ability and assertiveness using three items for each facet (see table 1). Traits were taken from prior psychometric studies (Koch et al., n.d.) and will be culturally adapted as described above. Participants will also complete two manipulation checks. A likeability measure to ensure Condition variable manipulation will be answered for all social targets (i.e., How much do you like or dislike this person/group? 0: Dislike them very much; 100: Like them very much). Second, an Inclusion of the Other in Self (IOS; Aron, Aron, &

³ See OSF page of the project for pre-registration, complete materials, a priori power analysis and pre-registered data analysis scripts: DOI 10.17605/OSF.IO/8F5S4.

Smollan, 1992) will be completed for all social targets involving groups (i.e., an ingroup, an outgroup, a person within an ingroup, and a person within an outgroup). IOS is a 7-point graphical scale measuring the extent to which another person or group is perceived to be close to the self. All items except for the IOS will be answered in a 100-point slider scale (see online supplementary materials). See below for pre-registered analyses and inclusion/exclusion criteria for these measures. Order within each scale and order of scale presentation will be randomised for each participant.

Table 1: Chosen adjectives for each dimension/facet.

Dimensions/facets	Adjectives
Horizontal: Morality	Honest Trustworthy Moral
Horizontal: Friendliness	Friendly Agreeable Sociable
Vertical: Ability	Skilled Competent Intelligent
Vertical: Assertiveness	Determined Self-assured Persistent

Finally, after the main data collection, participants will complete a short demographic questionnaire. Every lab will be able to add any additional measures they see fit after this demographic questionnaire to be used independently by them.

All labs will be instructed to adapt materials in a socially neutral way, for instance avoiding terms heavily linked to specific ethnic, socioeconomic or political groups that could be deemed to be insulting or slurs.

Pre Registered Analysis

Outlier Identification and Inclusion/Exclusion Criteria

Inclusion criteria include: being of legal age according to local legislation and being native in the language used for data collection. Participant exclusion criteria will include not describing social targets following experimental instructions (e.g., not describing a target group in the in-group condition, negatively describing a social target in the like condition...), completing the whole survey at extreme times ($3SD \pm \text{Mean Completion time}$) and failing at least one of the manipulation checks. The likeability manipulation check will be considered failed if participants answer below 60 or above 40 in a 100-point slider scale (for the like and dislike condition respectively). Similarly, the IOS attention check will be considered failed if participants answer below 5 or above 3 for the ingroup and outgroup conditions, respectively.

Given language barriers, each lab will be expected to conduct data validation of open-ended descriptions to ensure data quality. This data quality will imply ensuring that each participant actually described a social target following instructions (e.g., actually described a single person in the target person condition, the outgroup is actually a group the person does not belong to and so on). We expect significant differences of warmth and sociability attributions across those conditions if the like/dislike manipulation was effective. No other inclusion or exclusion criteria will be considered. We will exclude participants according to these criteria before any data analysis.

Pre-registered Data Analysis

Main data analysis will contrast predictions by all considered models of person perception (i.e., bi-dimensional “simple” and “complex” models, tri-dimensional “simple” and

“complex”, and control models, see figs 1 – 5). The bi-dimensional “simple” model fits items in two dimensions (i.e., “horizontal” and “vertical” dimensions) with no distinction between facets within each dimension (see fig 1). The bi-dimensional “complex” model also fits the “horizontal” and “vertical” dimensions, but distinguishes the horizontal dimension between Friendliness and Morality facets, and the vertical dimension is divided into Assertiveness and Ability facets (see fig 2). The tri-dimensional “simple” models distinguish three equally important: “horizontal”, “vertical” and “moral” dimensions. For this model, morality items are expected to load into a single dimension, as will friendliness and vertical dimension items. No facets within dimensions are considered in this model (see fig 3). Finally, the tri-dimensional “complex” model closely follows the tri-dimensional “simple” model but distinguishes Assertiveness and Ability facets within the vertical dimension (see fig 4). We assume that all considered traits are independent from one another, thereby not fitting cross-correlations between traits or between factors. All models will be compared to a theoretically irrelevant, uni-dimensional control model (fig 5).

We will fit one CFA reflecting each theorised model (figs 1 – 5) using the lavaan R package (Rosseel, 2012) to compare fit (RMSEA, CFI, TLI, and SRMS) and generalizability (AIC and BIC) indexes between them to determine which model better fits data. To ensure the interpretability of observed differences between goodness-of-fit indexes, we will only interpret observed fit differences as conclusive if their absolute value is $\Delta\text{RMSEA} > 0.07$, as determined by our power analyses. Any smaller observed fit difference between models will be deemed inconclusive due to inadequate power. See below lavaan notation for all models.

```
# Bi-dimensional “simple” model:
bifactor.simple <- "warmth =~ mor_1 + mor_2 + mor_3 + fri_1 + fri_2 + fri_3
compet =~ ass_1 + ass_2 + ass_3 + abi_1 + abi_2 + abi_3"
```

<pre># Bi-dimensional "complex" model: bifactor.complex <- "warmth =~ morality + friendly compet =~ assertive + ability morality =~ mor_1 + mor_2 + mor_3 friendly =~ fri_1 + fri_2 + fri_3 assertive =~ ass_1 + ass_2 + ass_3 ability =~ abi_1 + abi_2 + abi_3"</pre>
<pre># Tri-dimensional "simple" model: trifactor.simple <- "warmth =~ fri_1 + fri_2 + fri_3 moral =~ mor_1 + mor_2 + mor_3 compet =~ abi_1 + abi_2 + abi_3 + ass_1 + ass_2 + ass_3"</pre>
<pre># Tri-dimensional "complex" model: trifactor.complex <- "warmth =~ fri_1 + fri_2 + fri_3 moral =~ mor_1 + mor_2 + mor_3 compet =~ assertive + ability assertive =~ ass_1 + ass_2 + ass_3 ability =~ abi_1 + abi_2 + abi_3"</pre>
<pre># Control model: Unifactor control <- "unifactor =~ fri_1 + fri_2 + fri_3 + mor_1 + mor_2 + mor_3 + abi_1 + abi_2 + abi_3 + ass_1 + ass_2 + ass_3"</pre>
<pre># Control model: 4D fourfactor <- "assertive =~ ass_1 + ass_2 + ass_3 ability =~ abi_1 + abi_2 + abi_3 warmth =~ fri_1 + fri_2 + fri_3 moral =~ mor_1 + mor_2 + mor_3"</pre>

Note that all models reflect a different, theoretically justified, factorial organisation of the same set of 12 items. In essence, we will confront five theoretically distinct possible factorial organisations of the same dataset to determine which better fits the empirically observed data. In this sense, our study follows best practices in CFA and SEM models, where theoretically distinct models are pitted against one another to determine which, if any, better fits the data (Crede & Harms, 2019; Groskurth, Bluemke, & Lechner, 2022).

Exploratory Data Analysis

Exploratory data analysis will be directed at determining whether fundamental dimensions of person perception change according to the considered social target. For this question, we will use a similar data analysis strategy to the main research question by fitting and comparing all considered models (figs 1 – 5) according to their intended social target. Therefore, it will be possible for us to empirically test claims made by adversarial collaborations (Abele et al., 2020; Koch et al., 2021) suggesting that the considered social target determines how many fundamental dimensions are used to evaluate the target. As above, we will compare fit (RMSEA, CFI, TLI, and SRMS), and generalizability (AIC, BIC) indexes to determine whether the preferred model significantly varies according to the considered social target. To ensure accurate modelling, we will compare CFAs only when sufficient observations, that is a minimum of $n = 229$, are observed for each social target.

Cross-cultural Comparisons

Proposed samples allow for direct testing of cross-cultural variations by collecting sufficient data in a large variety of cultural contexts across the globe. Hence, we will run a Multi-Level Confirmatory Factor Analysis (MLCFA) to assess whether the obtained factorial structure varies across cultural contexts. Exploratory analyses will be run using individual countries, as well as using country-level data such as the Gini Index, Gender Inequality Index, Subjective Well-Being, Multidimensional Poverty Index, Cultural Distance, GDP per capita, collectivism-individualism, religiosity, Human Development Index, percentage of Urban Population, and Life Expectancy at birth as level 2 variables. As these models require a large amount of data collected in each country, the decision of which variables, including as level 2 predictors, will largely depend on the final sample collected in each country and the availability of the required data. Hence, they will be purely exploratory in nature.

Pre Published Manuscript DO NOT REPRODUCE

References

- Abele, A. E., Ellemers, N., Fiske, S. T., Koch, A., & Yzerbyt, V. (2020). Navigating the Social World: Toward an Integrated Framework for Evaluating Self, Individuals, and Groups. *Psychological Review*, 128(2), 290–314. <https://doi.org/10.1037/rev0000262>
- Abele, A. E., Hauke, N., Peters, K., Louvet, E., Szymkow, A., & Duan, Y. (2016). Facets of the fundamental content dimensions: Agency with competence and assertiveness-communion with warmth and morality. *Frontiers in Psychology*, 7(November), 1–17. <https://doi.org/10.3389/fpsyg.2016.01810>
- Abele, A. E., & Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of Personality and Social Psychology*, 93, 751–763. <https://doi.org/10.1037/0022-3514.93.5.751>
- Abele, A. E., & Wojciszke, B. (2014). Communal and agentic content in social cognition: A dual perspective model. In *Advances in Experimental Social Psychology* (1st ed., Vol. 50). <https://doi.org/10.1016/B978-0-12-800284-1.00004-7>
- Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of Other in the Self Scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, 63(4), 596–612. <https://doi.org/10.1037/0022-3514.63.4.596>
- Barbosa, S., Carmona-Diaz, G. M., Rojas, D. S., Pineda-Molano, J. G., Baez, R. S., Ribero, E. E., & Jiménez-Leal, W. (n.d.). Confronting stereotype bi-dimensional and tri-dimensions stereotype models in regional and country groups. *In Preparation*.
- Brambilla, M., & Leach, C. W. (2014). On the Importance of Being Moral: the Distinctive Role of morality in Social Judgment. *Social Cognition*, 32, 397–408. Retrieved from <internal-pdf://1.127.196.16/Brambilla, 2014, On the Importance of Being Mo.pdf>
- Brambilla, M., Rusconi, P., Sacchi, S., & Cherubini, P. (2011). Looking for honesty: The primary role of morality (vs. sociability and competence) in information gathering. *European Journal of Social Psychology*, 41, 135–143. <https://doi.org/10.1002/ejsp.744>
- Brambilla, M., Sacchi, S., Rusconi, P., & Goodwin, G. P. (2021). The primacy of morality in impression development: Theory, research, and future directions. In *Advances in Experimental Social Psychology* (Vol. 64, pp. 187–262). <https://doi.org/10.1016/bs.aesp.2021.03.001>
- Carrier, A., Louvet, E., Chauvin, B., & Rohmer, O. (2014). The primacy of agency over competence in status perception. *Social Psychology*, 45(5), 347–356. <https://doi.org/10.1027/1864-9335/a000176>
- Crede, M., & Harms, P. (2019). Questionable research practices when using confirmatory factor analysis. *Journal of Managerial Psychology*, 34(1), 18–30. <https://doi.org/10.1108/JMP-06-2018-0272>
- Critcher, C. R., Helzer, E. G., & Tannenbaum, D. (2020). Moral character evaluation: Testing another's moral-cognitive machinery. *Journal of Experimental Social Psychology*, 87(September 2019), 103906. <https://doi.org/10.1016/j.jesp.2019.103906>

- Cuddy, A. J., Fiske, S. T., & Glick, P. (2008). Warmth and Competence as Universal Dimensions of Social Perception: The Stereotype Content Model and the BIAS Map. *Advances in Experimental Social Psychology*, 40, 61–149.
[https://doi.org/10.1016/s0065-2601\(07\)00002-0](https://doi.org/10.1016/s0065-2601(07)00002-0)
- Ellemers, N., & van den Bos, K. (2012). Morality in Groups: On the Social-Regulatory Functions of Right and Wrong. *Social and Personality Psychology Compass*, 6(12), 878–889. <https://doi.org/10.1111/spc3.12001>
- Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: warmth and competence. *Trends in Cognitive Sciences*, 11, 77–83.
<https://doi.org/10.1016/j.tics.2006.11.005>
- Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2018). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. In *Social Cognition: Selected Works of Susan Fiske*.
<https://doi.org/10.4324/9781315187280>
- Fiske, S. T., Nicolas, G., & Bai, X. (2021). The stereotype content model: How we make sense of individuals and groups. In P. A. M. Van Lange, E. T. Higgins, & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (3rd ed., pp. 392–410). The Guilford Press.
- Friehe, M.-T., Kotzur, P. F., Böttcher, J., Zöller, A.-K. C., Lüttmer, T., Wagner, U., Asbrock, F., & Van Zalk, M. H. W. (2022). Examining the Structural Validity of Stereotype Content Scales – A Preregistered Re-Analysis of Published Data and Discussion of Possible Future Directions. *International Review of Social Psychology*, 35(1), 1.
<https://doi.org/10.5334/irsp.613>
- Goodwin, G. P., Piazza, J., & Rozin, P. (2014). Moral character predominates in person perception and evaluation. *Journal of Personality and Social Psychology*, 106, 148–168.
<https://doi.org/10.1037/a0034726>
- Groskurth, K., Bluemke, M., & Lechner, C. M. (2022). *Why We Need to Abandon Fixed Cutoffs for Goodness-of-Fit Indices: A Comprehensive Simulation and Possible Solutions*.
<https://doi.org/https://doi.org/10.31234/osf.io/5qag3>
- Hartley, A. G., Furr, R. M., Helzer, E. G., Jayawickreme, E., Velasquez, K. R., & Fleenor, W. (2016). Morality's Centrality to Liking, Respecting, and Understanding Others. *Social Psychological and Personality Science*, 7, 648–657.
<https://doi.org/10.1177/1948550616655359>
- Hester, N., & Hehman, E. (2020). Dress is a Fundamental Component of Person Perception. *Personality and Social Psychology Review*, 5(3), 248–253.
- Jobst, L. J., Bader, M., & Moshagen, M. (2021). A tutorial on assessing statistical power and determining sample size for structural equation models. *Psychological Methods*, No Pagination Specified-No Pagination Specified. <https://doi.org/10.1037/met0000423>
- Koch, A., Imhoff, R., Dotsch, R., Unkelbach, C., & Alves, H. (2016). The ABC of Stereotypes About Groups: Agency/Socioeconomic Success, Conservative-Progressive Beliefs, and

- Communion. *Journal of Personality and Social Psychology*, 110(5), 675–709.
<https://doi.org/10.1037/pspa0000046>
- Koch, A., Smith, A., Fiske, S. T., Abele, A. E., Ellemers, N., & Yzerbyt, V. Y. (n.d.). Validating a brief measure of four facets of social evaluation. *Under Review*.
- Koch, A., Yzerbyt, V., Abele, A., Ellemers, N., & Fiske, S. T. (2021). Social evaluation: Comparing models across interpersonal, intragroup, intergroup, several-group, and many-group contexts. In *Advances in Experimental Social Psychology* (1st ed., Vol. 63).
<https://doi.org/10.1016/bs.aesp.2020.11.001>
- Kosakowska-Berezecka, N., Bosson, J. K., Jurek, P., Besta, T., Olech, M., Vandello, J. A., ... Żadkowska, M. (2023). Gendered Self-Views Across 62 Countries: A Test of Competing Models. *Social Psychological and Personality Science*, 14(7), 808–824.
<https://doi.org/10.1177/19485506221129687>
- Landy, J. F., Piazza, J., & Goodwin, G. P. (2016). When It's Bad to Be Friendly and Smart: The Desirability of Sociability and Competence Depends on Morality. *Personality and Social Psychology Bulletin*, 42, 1–19. <https://doi.org/http://dx.doi.org/10.1177/0146167216655984>
- Leach, C. W., Ellemers, N., & Barreto, M. (2007). Group Virtue: The Importance of Morality (vs. Competence and Sociability) in the Positive Evaluation of In-Groups. *Journal of Personality and Social Psychology*, 93(2), 234–249.
<https://doi.org/10.1037/0022-3514.93.2.234>
- López-Rodríguez, L., Cuadrado, I., & Navas, M. (2013). Aplicación extendida del Modelo del Contenido de los Estereotipos (MCE) hacia tres grupos de inmigrantes en España. *Estudios de Psicología*, 34(2), 197–208. <https://doi.org/10.1174/021093913806751375>
- Naumann, L. P., Vazire, S., Rentfrow, P. J., & Gosling, S. D. (2009). Personality judgments based on physical appearance. *Personality and Social Psychology Bulletin*, 35, 1661–1671.
<https://doi.org/10.1177/0146167209346309>
- Newman, G. E., Bloom, P., & Knobe, J. (2014). Value judgments and the true self. *Personality and Social Psychology Bulletin*, 40, 203–216. <https://doi.org/10.1177/0146167213508791>
- Nicolas, G., Bai, X., & Fiske, S. T. (2022). A Spontaneous Stereotype Content Model : Taxonomy , Properties , and Prediction. *Journal of Personality and Social Psychology*.
<https://doi.org/https://doi.org/10.1037/pspa0000312>
- Pizarro, D. A., & Tannenbaum, D. (2011). Bringing Character Back: How the Motivation to Evaluate Character Influences Judgments of Moral Blame. In M. S. Milkulincer P. (Ed.), *The Social Psychology of Morality: Exploring the Causes of Good and Evil*. Retrieved from internal-pdf://196.4.165.224/Pizarro, 2011, Bringing character back.pdf
- Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48(2 SE-Articles), 1–36. <https://doi.org/10.18637/jss.v048.i02>
- Sayans-Jiménez, P., Cuadrado, I., Rojas, A. J., & Barrada, J. R. (2017). Extracting the evaluations of stereotypes: Bi-factor model of the stereotype content structure. *Frontiers in Psychology*, 8(1692), 1–12. <https://doi.org/10.3389/fpsyg.2017.01692>

- Stanley, M. L., Bedrov, A., Cabeza, R., Brigard, F. De, Stanley, M. L., Bedrov, A., ... Brigard, F. De. (2019). The centrality of remembered moral and immoral actions in constructing personal identity. *Memory*, 28(2).
<https://doi.org/https://doi.org/10.1080/09658211.2019.1708952>
- Stanley, M. L., Henne, P., Brigard, F. De, & Stanley, M. L. (2018). Remembering moral and immoral actions in constructing the self. *Memory and Cognition*, 47, 441–454.
- Strohminger, N., Knobe, J., & Newman, G. (2017). The True Self: A Psychological Concept Distinct From the Self. *Perspectives on Psychological Science*, 12(4).
<https://doi.org/10.1177/1745691616689495>
- Uhlmann, E., Pizarro, D. A., & Diermeier, D. (2015). A Person-Centered Approach to Moral Judgment. *Perspectives on Psychological Science*, 10, 72–81.
<https://doi.org/10.1177/1745691614556679>
- Leach, C. W., Ellemers, N., & Barreto, M. (2007). Group Virtue: The Importance of Morality (vs. Competence and Sociability) in the Positive Evaluation of In-Groups. *Journal of Personality and Social Psychology*, 93(2), 234–249.
<https://doi.org/10.1037/0022-3514.93.2.234>
- Yzerbyt, V. Y. (2018). The dimensional compensation model: Reality and strategic constraints on warmth and competence in intergroup perceptions. *Agency and Communion in Social Psychology*, 126–141. <https://doi.org/10.4324/9780203703663>
- Yzerbyt, V. Y., Kervyn, N., & Judd, C. M. (2008). Compensation versus halo: the unique relations between the fundamental dimensions of social judgment. *Personality and Social Psychology Bulletin*, 34, 1110–1123. <https://doi.org/10.1177/0146167208318602>