UNIVERSITY of CALIFORNIA PRESS | Collabra: Psychology

Barbedor, J., Scheider, J., Yzerbyt, V., & Abele-Brehm, A. (2024). A Novel Approach to the Evaluation of Groups: Type of Group and Facet of Evaluation Matter. *Collabra: Psychology*, *10*(1). https://doi.org/10.1525/collabra.124471

Social Psychology

A Novel Approach to the Evaluation of Groups: Type of Group and Facet of Evaluation Matter

Julien Barbedor[1][a], Jelena Scheider[2][b], Vincent Yzerbyt[1][c], Andrea Abele-Brehm[2][d]

[1] Institut de Recherche en Sciences Psychologiques, Université catholique de Louvain, Louvain-la-Neuve, Belgium, [2] Department of Psychology, Friedrich-Alexander-Universität Erlangen-Nürnberg

Keywords: Social evaluation, Social perception, Stereotypes, Type of groups

https://doi.org/10.1525/collabra.124471

Collabra: Psychology

Vol. 10, Issue 1, 2024

In three studies, we examined a novel approach to the social evaluation of groups whereby we combined a typology of groups with a recent model of the two fundamental dimensions and four facets of social evaluation. We investigated two main questions. First, whereas previous research has proven the usefulness of the two-dimension/four-facet model for individuals, we here tested it for groups. Second, whereas previous research has not systematically distinguished between different types of groups, we here applied a typology of groups that distinguishes between intimacy groups, task groups, and social categories. Furthermore, we hypothesized that these dimensions and facets would manifest differently across various group types. We conducted studies in three countries and analyzed judgments of the three main types of groups, resulting in a standard list of 20 traits. The data confirmed our two-dimension/four-facet model while showing variations of the structure across different types of groups. These results, as well as the facets ratings, highlight the relevance of distinguishing the types of groups in social evaluation.

People spend most of their time in groups. They live in families, they meet with their friends, they are members of sports teams, they form part of work groups or belong to social, political, religious, or even virtual groups. People often talk about groups, both groups to whom they belong and groups that they know without being member. They read and hear about groups, most often in the media, for instance political parties or religious communities. Because groups are such an essential part of life, people are keen on evaluating them. Do they like them? Do they trust them? Is a specific group competent or not? These questions are but a few examples of the basic goals that underlie social evaluation (Abele & Wojciszke, 2007; Fiske et al., 2007; Yzerbyt & Demoulin, 2010).

We are here concerned with a novel approach to the evaluation of groups. We suggest combining two lines of research that have not been connected heretofore. On the one hand, we build on research investigating the differences between various types of groups (Lewin, 1948). As shown by Lickel and colleagues (2000, 2001), groups differ in some systematic ways. In all likelihood, evaluation should differ depending on these different types of groups. On the other hand, we follow the work on the fundamental dimensions or "Big Two" of social evaluation (for reviews, see Abele et al., 2021; Koch et al., 2021). As many scholars pointed out, evaluations of the self, of others, and of groups rest on two fundamental dimensions that have recently been labeled vertical (also agency or competence) and horizontal (also communion or warmth), each of these comprising two facets (vertical: assertiveness and ability; horizontal: friendliness and morality). We argue that combining these two lines of research is useful both with respect to the understanding of the evaluation of different kinds of groups and with respect to the understanding of the fundamental dimensions of social judgment.

We report on three studies that examine this novel approach. We aimed to investigate two main research questions. First, we tested the hypothesis that the dimensions and facets of social evaluation can be measured reliably with respect to groups. Second, we tested predictions re-

a Correspondence concerning this article should be addressed to Julien Barbedor, Institut de Recherche en Sciences Psychologiques, Université catholique de Louvain, Place du Cardinal Mercier, 1348, Louvain-la-Neuve, Belgium. email: julien.barbedor@uclouvain.be

Shared first authorship

b Shared first authorship

c Shared last authorship

d Shared last authorship

garding how perceivers evaluate different types of groups on these facets.

Types of Groups

Social psychology offers many definitions of social groups (Augoustinos & Walker, 2001; Sherif & Sherif, 1956). Following Brown (1998), a group exists if two or more people perceive themselves and are perceived by at least one observer as a group. Social groups can vary on a large number of features and people can identify with and perceive groups for a host of reasons (Brehm et al., 1999; Yzerbyt et al., 2004). In the literature, groups have often been distinguished with respect to specific properties, such as the intimacy of a group in primary and secondary groups (Cooley, 1909) or the size of the group (La Macchia et al., 2016).

Lewin (1948) was among the first authors to systematically distinguish between different types of groups. He distinguished between social categories and dynamic groups. According to his view, social categories are groups that identify individuals with certain characteristics, for instance, nationality, occupation, or hair color. They are similar with respect to such a characteristic. Dynamic groups, in contrast, are interdependent and change over time. Similarity is not the main characteristic. "Intimate groups" (e.g., friends, family) and "task groups" (e.g., people working together on specific tasks) are examples of dynamic groups. In a seminal effort, Lickel et al. (2000) provided empirical support for this typology. These authors examined how people intuitively evaluated and sorted 40 groups into different categories. They identified four distinct types, i.e., intimacy groups, task groups, social categories, and loose associations. These types vary systematically in several ways: Duration, Interaction, Shared Goals and Outcomes, Permeability, Meaning, Similarity, and Entitativity (Campbell, 1958; being perceived as "group-like" or an entity). Intimacy groups such as friends or family, come across as highly entitative (Denson et al., 2006; Hamilton et al., 2002; Lickel et al., 2000, 2001), as being important (Carnes et al., 2015; Denson et al., 2006; Lickel et al., 2000, 2001), and as having a limited number of members who interact frequently. Task groups such as the crew of an airplane, the actors in a play, or the members of a project team, are characterized by common goals and intended outcomes (Lickel et al., 2000, 2001). They are perceived as less of an entity than intimacy groups. The third type of groups is social categories. Social categories are mainly defined by one specific characteristic that is superimposed on a number of people sharing this characteristic (e.g., being a man vs. being a woman; being old vs. being young). People belonging to one social category do not necessarily interact with each other. In order to perceive them as a social group it is sufficient that they share the characteristic. Compared to intimacy and task groups, social categories are seen as less entitative (but see Castano et al., 2003; Yzerbyt et al., 2000). Finally, loose associations are, for instance, people who wait in a line together or people who share the same neighborhood. Those groups are similar to simple aggregates of people and are not seen as entitative (Denson et al., 2006; Lickel et al., 2000, 2001).

This typology of groups has been supported by studies using implicit tasks (Sherman et al., 2002). Even 5- to 6-year old children already distinguish those different types of groups (Plötner et al., 2016). Moreover, studies showed that observers link those types of groups to differential beliefs about morality principles (Carnes et al., 2015) and relational styles (Lickel et al., 2001, 2006), but also that those types of groups serve different functional needs (Crawford & Salaman, 2012).

The "Big Two" of social evaluation and their facets

A large body of research shows that social evaluation can be mapped onto two dimensions. People perceive and evaluate themselves and others on the so-called Big Two, namely Warmth or Communion on the one hand and Competence or Agency on the other (Abele & Wojciszke, 2014; Fiske, 2015; Fiske et al., 2002; Judd et al., 2005; Yzerbyt, 2016; Yzerbyt et al., 2005). For the sake of unifying the impressive number of studies that deal with these two dimensions but use different labels, recent work suggests using the labels Horizontal dimension vs. Vertical dimension (for a review, see Abele et al., 2021). The Horizontal dimension subsumes traits like trustworthy, honest, or sincere but also such characteristics as friendly, likeable, or sociable. From a functional perspective, the Horizontal dimension depicts the challenge of recognizing a target's intention while, from an evolutionary perspective, it evaluates the extent to which a target is able to form bonds and maintain social relationships (Abele et al., 2021; Abele & Wojciszke, 2014; Fiske, 2018). As to the Vertical dimension, it denotes the extent to which people are able to interact on the basis of those intentions (Fiske et al., 2007) and is linked to aspects of power and dominance (Abele & Wojciszke, 2014; Yzerbyt, 2016). Related traits predict goal-achievement and task-functioning, like competence, efficiency, or ability as well as assertiveness, determination, and self-assurance (Abele & Wojciszke, 2014).

Recent developments in social evaluation research propose that the two basic dimensions can be differentiated into two facets each (Abele et al., 2008). Whereas the Horizontal dimension entails a *friendliness* facet and a *morality* facet, the Vertical dimension comprises an *assertiveness* facet and an *ability* facet, (Abele et al., 2008, 2016, 2021). To achieve goals, not only intellectual and practical skills such as cleverness are important (ability), but also the motivation and volition to pursue these goals (assertiveness). To form social relationships, it is important that the person comes across as likable and has the skill to be agreeable (friendliness), but also that he/she has a benevolent underlying motivation, and is perceived as trustworthy (morality) (Wojciszke & Abele, 2018).

Interestingly, similar differentiations emerge in other research. In accordance with the postulated horizontal facets, several empirical efforts showed that the evaluation of morality can differ from a friendliness evaluation in a group context (Brambilla et al., 2012; Ellemers, 2017; Elle-

mers et al., 2008; Leach et al., 2007). With respect to the vertical facets, studies highlighted the importance of distinguishing between assertiveness and competence as they are differentially related to status (Carrier et al., 2014; Louvet et al., 2019; Yzerbyt et al., 2022).

This two-dimension/four-facet structure was empirically validated with respect to individuals (Abele et al., 2016; Abele & Hauke, 2019). Moreover, the distinction of – at least, some of – the facets per Big Two dimension has already proven fruitful in a number of studies. It has been shown that the assertiveness facet of the Vertical dimension is more related to self-esteem and also to self-efficacy than the ability facet (Abele, 2022; Abele & Hauke, 2019). Assertiveness self-perception is more related to economic success than ability self-perception (Miraucourt et al., 2022). Not only is status more related to assertiveness than ability (Louvet et al., 2019) but this differential relation materializes more readily in the eyes of low social dominant or low-status perceivers (Yzerbyt et al., 2022). Others are liked and respected more, when they are seen as high in ability, but not high in assertiveness (Abele & Hauke, 2019). People gossip more about others' low warmth, low morality, and low competence, but – in contrast – more about others' high assertiveness (Martinescu et al., 2022). Group identification is more related to morality than to friendliness (also called sociability; Brambilla et al., 2021; Leach et al., 2007). Finally, whereas the general valence – the evaluation of traits irrespective of a specific target – of the facets of the Horizontal dimension is roughly the same, and the general valence of the ability facet of the Vertical dimension is similar to that of the Horizontal facets, the assertiveness facet is generally rated less positively – irrespective of the fact that it is highly related to self-esteem (Abele, 2022).

The four facets in the evaluation of different types of groups

Combining the above group typology with the two-dimension/four-facet approach, we wanted to extend previous research in several important ways. Previous work on the dimensions and facets was concerned with individuals and less so, if at all, with groups. Moreover, this work did not cover all four horizontal and vertical facets distinguished here, but rather concentrated on a subsample (Leach et al., 2007; Louvet et al., 2019; Yzerbyt et al., 2022). Finally, groups were studied on the basis of specific characteristics, but not in light of a systematic typology (Burkley et al., 2017; Cuddy et al., 2009; Imhoff et al., 2018; Nicolas, Fiske, et al., 2022). Indeed, using the typology of groups proposed by Lickel et al. (2000, 2001), social categories are unquestionably the groups studied most in social evaluation research. Even research attempting to measure the evaluation of a representative sample of groups did not capture intimacy groups and task groups or loose associations (Koch et al., 2016, 2020; Nicolas, Bai, et al., 2022; Nicolas, Fiske, et al., 2022). To our knowledge, the only convincing effort combining the recent facet approach and the evaluation of groups is the one by Nicolas, Bai, et al. (2022). These authors asked participants to spontaneously evaluate groups (social categories only) and to generate any traits or characteristics that would describe the target. Cluster analyses revealed that the most prevalent dimensions used to describe groups were indeed the four facets. This result emphasizes the relevance of each of the facets and their primacy when it comes to social evaluation. However, this effort investigated spontaneous stereotypes relying on self-generated traits, preventing (1) the development of a set of items to capture the facet constructs, and (2) any factorial analyses to see how the facet structure applies to groups. Moreover, these results only concern social categories as targets and as such elude some important aspects of group evaluation.

The present research

The present research had two main goals. First, we aimed to explore and test whether the two-dimension/four-facet model could apply to the evaluation of groups. We hypothesized that the two-dimension/four-facet model observed in the evaluation of individuals should also operate in the evaluation of groups and that this model is superior to a two-dimension model. We conducted our studies in three different countries (Belgium, Germany, and Britain) with three different languages (French, English, and German) to test the generalizability of our findings.

Second, we analyzed the evaluation of three different types of groups explained above, e.g. social categories, intimacy groups, and task groups. We did not include the fourth type of groups distinguished by Lickel et al. (2000), namely loose associations, as these hardly come across as "group like" (Denson et al., 2006; Lickel et al., 2000; Spencer-Rodgers et al., 2007). We hypothesized that intimacy groups are rated higher on the Horizontal dimension, e.g., friendliness and morality, than task groups and social categories. Task groups should be higher on the Vertical dimension, e.g. assertiveness and ability, than intimacy groups and social categories. We also studied in a more exploratory fashion how the facet ratings differ both between and within group types.

We operationalized intimacy groups by the most representative examples revealed in the literature, namely family (Study 1 and Study 3) and friends (Study 1). We operationalized task groups as people who work together (Study 2 and Study 3). Social categories can be operationalized by many characteristics as outlined above. We here relied on previous research that analyzed which groups would come into mind when participants are asked to spontaneously generate several groups (Koch et al., 2016). By using this data, we tried to gather a representative sample of social categories (Study 1). We also specifically analyzed one social category, to which all study participants belonged, namely nationality (Study 3).

In all three studies, we also assessed further relevant features of groups, to make sure that we adequately operationalized the three types of groups. These features were entitativity, i.e., the degree to which an aggregate of people is perceived as an "entity", and a series of "relational" properties, namely the degree to which a person feels that this group is important for him/her, how similar a person feels to a group, and how much a person identifies with a group

(Denson et al., 2006; Lickel et al., 2000, 2001; Yzerbyt et al., 1997, 2000). These ratings checked for the adequacy of our operationalizations of group types. We predicted that entitativity ratings would be highest for intimacy groups and lowest for social categories. Relational features should again be highest for intimacy groups and lowest for social categories.

In sum, the three studies reported here aimed at enhancing our understanding of the two-dimension/four-facet model of social evaluation as well as our understanding of the evaluation of different types of groups.

Study 1

In Study 1, we presented participants with intimacy groups and social categories and measured their evaluations on items corresponding to one of the four facets of ability, assertiveness, friendliness, and morality. We also assessed perceived entitativity of the groups and perceivers' relation to the groups. We conducted the study in Belgium (French, Study 1a) and Germany (German, Study 1b).

Regarding intimacy groups, we asked our participants to either think of their "family" or to think of their group of "friends". The operationalization of social categories used Koch et al.'s (2016) findings. These authors asked their participants to list the social groups that spontaneously came to mind. We examined the list compiled in Germany (Koch et al., 2016, Table 5) and selected eight groups, four of the most frequently listed groups (students, athletes, unemployed, politicians), and four of the less frequently listed groups (teachers, vegetarians, musicians, and scientists). By basing our selection of social categories on previous work on the spontaneous generation of social categories, our materials ensured that the groups retained was representative of the huge number of possible social categories.

Participants thus evaluated 10 groups, two intimacy groups, and eight social categories. Because evaluating this number of groups with almost 30 items for each would have been exhausting to participants, we decided that each participant would evaluate only three groups, one intimate group, and two social categories. Participants were randomly assigned to the different combinations of groups.

Study 1a: Belgium

Method

Participants

We collected our data online. We distributed an online Qualtrics questionnaire via Facebook ads in Wallonia (French-speaking region of Belgium) and in Brussels (mostly French-speaking area). We excluded all participants who did not speak French as their native language or did not show any variance in their responses (same response to the entire scale). The sample comprised $N = 411$ participants ($n = 97$ male, $n = 313$ female, $n = 1$ other). The mean age was $M = 41.18$ years ($SD = 9.69$). 87% of participants held a university degree, and 13% had at least a secondary degree. The majority of the sample was employed ($n = 310$, 75.4%) or student ($n = 40$, 9.7%).

Procedure

After providing instructions and securing consent, we asked participants to evaluate three different groups. Each group was described shortly and followed by the various rating scales. Participants always rated three groups in random order: one intimacy group (randomly selected from "family" or "friends") and two social categories (randomly selected from the eight groups described above). We further randomized the order of presentation of the three groups.

Measures

Facet ratings. Following a short description of the group, participants had to report how they perceived the group on a series of traits on a 7-point scale ranging from 1 = not at all to 7 = very much. We randomized trait order. We took the traits from previous studies on the four facets (Abele et al., 2016; Abele & Hauke, 2019) as well as from other studies on group evaluation (e.g. Fiske, 2015; Yzerbyt et al., 2005; Yzerbyt & Cambon, 2017). We aimed to build scales with five items per facet that would be reliable in both German and French. To limit language discrepancies and maximize construct validity for each facet, we included more than five items per facet. The final item pool consisted of 27 items. One filler item (humorous) served as an example and appeared systematically at the beginning of each group rating phase (for the full item set, see Appendix A1).

Feature ratings and demographics. Following the facet ratings, participants rated their relation towards the group on three items, "importance" ("how important is this group for you"), perceived similarity ("how similar to yourself do you perceive the group to be"), and identification ("how much do you identify with this group") on 7-point rating scales ranging from "not at all" to "very much". Next, we asked participants to rate the group's entitativity. Building on Lickel et al. (2000) and Yzerbyt et al. (2001), we presented participants with a short explanation[1] and asked them to indicate the degree to which they perceived the group as a "real group" on a 7-point scale from 1 "not at all" to 7 "very much". Regarding the intimacy group (either family or friends), participants also indicated the rough number of group members. Finally, participants provided

1 The explanation read as follows: "All social groups are made up of several individuals. At the same time, groups of individuals vary in terms of the strength of the bond between their members, ranging from very loosely connected (e.g. a line of people waiting for the bus) to very tightly connected (people living in the same house)."

their gender, age, mother language, nationality, highest degree, and occupation.

Study 1b: Germany

We collected our data online. German participants received an online invitation and questionnaire (via the platform www.soscisurvey.de). We excluded participants who did not speak German as their mother language and participants who showed no variance in their responses. The German sample comprised $N = 394$ participants ($n = 201$ male, $n = 185$ female, $n = 8$ other; age: $M = 30.86$ years, $SD = 10.09$). 47% of participants held a university degree, and 40% had at least a secondary degree. Most participants were employed ($n = 207$, 52.5%), or students ($n = 147$, 37.3%). The procedure, measures, and analyses were the same as in Study 1a (Belgium).

Analyses

A first step was to construct reliable facet scales. We conducted principal component analyses (PCAs) on all the items used to measure the facets to reduce cross-loading items or items with low communalities. Based on these analyses, we selected 20 items to be included into the facet scales and tested their reliabilities. Using this 20-item selection, we relied on confirmatory factor analyses (CFAs) to test the fit of several models (one-dimension model, two-dimension model, four-facet model, and two-dimension/four-facet model). These analyses allowed testing our first research hypothesis, namely that the two-dimension/four-facet model holds for group evaluation, and that this model is superior to a two-dimension model.

In a second step, we focused on the types of groups. We first tested the differences in entitativity, identification, importance, and similarity ratings between the two types of groups to ascertain that intimacy groups are rated as more entitative, and the relation towards this type of group is more positive than towards social categories. We then applied the previously constructed facet scales and examined how people evaluate the two types of groups on the facets. These analyses tested our second research question, namely that perceivers evaluate differently the different types of groups.

The data and code of the analyses for all studies can be accessed via https://osf.io/3srq5/ on the Open Science Framework. We did not preregister these studies.

Results of Studies 1a and 1b

Scale construction

First, we conducted study-wise principal component analyses for the two types of groups with all items included. These analyses served as a first screening for items that showed high cross-loadings (>|.40|) and/or low communalities ($h^2 < |.50|$; (for analyses, see Supplementary material tables S2.1- S2.4). Based on these analyses, we retained five items per facet with low cross-loadings and high communalities in the respective language[2]. Table S2.10 shows eigenvalues for this and all remaining studies.

Confirmatory factor analyses. The selected 20 items were included into CFAs. The test of our theoretical two-dimension/four-facet model confirmed that the fit was acceptable in both languages (Table 1). Fit indices of the two-dimension/four-facet model were always higher than cut-off values (Hu & Bentler, 1999).

Within each sample, we compared the fit of the different models. The two-dimension model always fared significantly better than the one-dimension model ($p < .001$). The four-facet and two-dimension/four-facet models never differed significantly from each other while both always performed better than the two-dimension models ($p > .05$). These findings support our hypotheses that the two-dimension/four-facet model also applies to the evaluation of groups and that the two-dimension/four-facet model is superior to the two-dimension model. Because the recent theoretical and empirical literature (for reviews, see Abele et al., 2021; Koch et al., 2021) provides ample reasons to prefer a two-dimension/four-facet model over a four-facet model, we decided to stick to the former in the present context.

We then performed invariance analyses between the samples of Studies 1a and 1b. Taking into account the hierarchical structure of our model, we found a satisfactory fit with respect to configural invariance ($\chi^2(330) = 3445.54$, $p < .001$, CFI =.92, RMSEA =.09, SRMR =.08) allowing us to conclude that the data from the two samples supported the two-dimension/four-facet structure. The test of metric invariance differed statistically from the configural invariance test ($\Delta\chi^2(18) = 304$, $p<.001$), preventing us from directly comparing the results between samples.

Facet scales. We constructed the facet scales by averaging the responses to the respective traits. The traits belonging to the scales as used in French, German, and English are displayed in the Appendix. The descriptive statistics by item (S3), type of group, and group (S4) can be found in supplementary materials. The reliabilities ranged from acceptable to excellent, with Cronbach's α ranging from .81 to .94. Table 2 shows the intercorrelations between the facets in the three types of groups.

Supporting our model, assertiveness always correlates the highest with ability and friendliness correlates the highest with morality in most cases. At the same time, ability also correlates highly with morality (and friendliness).

2 It should be noted that not all items have direct correspondence in both languages. This is due to the fact that we tried to have similar meaning of the scales and this was in some cases only possible by selecting items that had no literal translation in the other language – a common problem in cross-lingual research (Ægisdóttir et al., 2008).

Table 1. Confirmatory Factor Analyses across studies

	χ2	*df*	*χ2/df*	*p*	AIC	CFI	RMSEA	SRMR
				Belgium (Study 1a)				
One-dimension model	6,063.39	170	35.70	< .001	72,604	.67	.17	.13
Two-dimension model	3,974.62	169	23.52	< .001	70,517	.79	.14	.11
Four-facet model	1,505.01	164	9.18	< .001	68,057	.93	.08	.10
Two-dimension/ four-facet model	1,505.03	165	9.12	< .001	68,055	.93	.08	.10
				Germany (Study 1b)				
One-dimension model	8,051.18	170	47.36	< .001	69,584	.62	.20	.17
Two-dimension model	4,338.08	169	25.67	< .001	65,872	.80	.14	.12
Four-facet model	1,912.86	164	11.66	< .001	63,457	.92	.09	.07
Two-dimension/ four-facet model	1,940.52	165	11.76	< .001	63,483	.91	.10	.08
				Belgium (Study 2a)				
One-dimension model	630.5	170	3.71	< .001	9119.78	.78	.13	.08
Two-dimension model	451.98	169	2.67	< .001	8943.26	.87	.11	.07
Four-facet model	318.91	164	1.94	< .001	8820.18	.93	.08	.06
Two-dimension/four-facet model	319.19	165	1.93	< .001	8818.47	.93	.08	.06
				Germany (Study 2b)				
One-dimension model	696.61	170	4.10	< .001	7205.42	.69	.16	.12
Two-dimension model	424	169	2.51	< .001	6934.81	.85	.11	.09
Four-facet model	354.44	164	2.16	< .001	6875.25	.89	.10	.08
Two-dimension/ four-facet model	356.42	165	2.16	< .001	6875.22	.89	.10	.08
				United Kingdom (Study 3)				
One-dimension model	2,006.37	170	11.80	< .001	32,065	.80	.14	.08
Two-dimension model	1,247.64	169	7.38	< .001	31,308	.88	.10	.06
Four-facet model	662.50	164	4.04	< .001	30,733	.95	.07	.05
Two-dimension/ four-facet model	663.48	165	4.02	< .001	30,732	.95	.07	.05

Note. AIC = Akaike Information Criterion; CFI = comparative fit index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual.

Group features

We tested whether the group types differed in terms of entitativity, as suggested by previous research (Lickel et al., 2000). We also examined the ratings of importance, similarity, and identification. For both Studies 1a and 1b, we conducted a mixed model with the type of group as a predictor and entitativity ratings as the criterion. We treated participants and groups as random intercepts. Confirming our hypotheses as well as previous findings, entitativity ratings were significantly higher for intimacy groups than for social categories (see Table 3), and intimacy groups were also rated significantly higher on identification, similarity, and importance compared to social categories. This pattern confirms that our operationalization of intimacy groups and social categories was successful.

Group types and facet evaluation

Figure 1 presents the means and standard errors of the facet scales for the two types of groups in both studies. For both studies, we conducted a 2 (type of group) x 4 (facet) mixed model. We treated the factors and their interactions as fixed effects varying within participants and implemented random intercepts for participants, groups, and items. We conducted all of the analyses with the *lme4* package in R and computed the effect sizes with the *r2glmm*

Table 2. Correlations between the facets across the studies

		Ability	Assertiveness	Friendliness	Morality
			Study 1a (Belgium)		
Intimacy groups	Ability	1.00	0.63***	0.64***	0.68***
	Assertiveness		1.00	0.49***	0.45***
	Friendliness			1.00	0.74***
	Morality				1.00
Social categories	Ability	1.00	0.52***	0.59***	0.75***
	Assertiveness		1.00	0.26***	0.31***
	Friendliness			1.00	0.65***
	Morality				1.00
			Study 1b (Germany)		
Intimacy groups	Ability	1.00	0.74***	0.59***	0.60***
	Assertiveness		1.00	0.37***	0.39***
	Friendliness			1.00	0.79***
	Morality				1.00
Social categories	Ability	1.00	0.66***	0.45***	0.69***
	Assertiveness		1.00	0.16***	0.39***
	Friendliness			1.00	0.76***
	Morality				1.00
			Study 2a (Belgium)		
Task group	Ability	1.00	0.76***	0.57***	0.75***
	Assertiveness		1.00	0.56***	0.71***
	Friendliness			1.00	0.78***
	Morality				1.00
			Study 2b (Germany)		
Task group	Ability	1.00	0.79***	0.53***	0.70***
	Assertiveness		1.00	0.37***	0.53***
	Friendliness			1.00	0.78***
	Morality				1.00
			Study 3 (United Kingdom)		
Intimacy group	Ability	1.00	0.65***	0.61***	0.67***
	Assertiveness		1.00	0.58***	0.55***
	Friendliness			1.00	0.78***
	Morality				1.00
Task group	Ability	1.00	0.72***	0.62***	0.76***
	Assertiveness		1.00	0.49***	0.60***
	Friendliness			1.00	0.76***
	Morality				1.00
Social category	Ability	1.00	0.69***	0.61***	0.73***
	Assertiveness		1.00	0.55***	0.60***
	Friendliness			1.00	0.62***
	Morality				1.00

Note. *** p<.001.

package using the *nsj* method. Due to multiple comparisons, we only report significant effects using a threshold of p <.01.

Facet ratings between types of groups. For Study 1a (Belgium), the analysis revealed a type of group by facet interaction, $F(3, 24218) = 658.29$, p <.001, showing higher ratings in friendliness (b = -1.17, $t(8)$ = -3.41, p =.009, r^2 =.040) and morality (b = -1.50, $t(8)$ = -4.37, p =.002, r^2 =.064) for intimacy groups than for social categories. Similarly, for Study 1b (Germany), the type of group by facet interaction was significant, $F(3, 23215) = 652.52$, p <.001, revealing higher ratings in friendliness (b = -1.46, $t(8)$ = -3.57, p

Table 3. Ratings on additional variables depending on the group type across the studies

	Intimacy group	Task- group	Social category	*t*-value	r^2
	M (SD)	*M (SD)*	*M (SD)*		
			Belgium (Study 1a)		
Entitativity	4.92_a (1.77)	NA	3.88_b (1.74)	-1.93	.067
Identification	4.96_a (1.46)	NA	2.61_b (1.75)	-5.75	.301
Similarity	4.81_a (1.44)	NA	2.85_b (1.72)	-4.64	.238
Importance	5.97_a (1.30)	NA	4.05_b (1.92)	-2.89	.201
			Germany (Study 1b)		
Entitativity	5.34_a (1.55)	NA	3.94_b (1.63)	-3.45	.141
Identification	5.53_a (1.43)	NA	2.79_b (1.83)	-4.16	.349
Similarity	5.14_a (1.45)	NA	3.06_b (1.74)	-3.55	.25
Importance	6.11_a (1.19)	NA	3.67_b (1.71)	-5.19	.343
			Belgium (Study 2a)		
Entitativity	NA	4.13 (1.59)	NA		
Identification	NA	3.84 (1.62)	NA		
Similarity	NA	3.95 (1.58)	NA		
Importance	NA	4.56 (1.64)	NA		
			Germany (Study 2b)		
Entitativity	NA	4.63 (1.64)	NA		
Identification	NA	4.04 (1.74)	NA		
Similarity	NA	4.34 (1.61)	NA		
Importance	NA	4.47 (1.75)	NA		
			United Kingdom (Study 3)		
Entitativity	5.60_a (1.21)	4.68_b (1.20)	3.48_c (1.21)	7.69; 10.00; 17.72	.088; .141; .34
Identification	6.01_a (1.40)	4.61_b (1.67)	4.53_b (1.54)	9.81; 10.39	.133; .12
Similarity	5.53_a (1.54)	4.43_b (1.59)	4.37_b (1.52)	7.84; 8.29	.076; .084
Importance	6.53_a (1.01)	4.63_b (1.77)	4.35_b (1.70)	13.15; 15.15	.203; .253

Note. Means with different suffixes differ at the 1% level (rows). T-tests and effect sizes in Study 3 are for intimacy versus task group, task group and social categories, and intimacy and social categories, respectively. NA indicates that the data were not available.

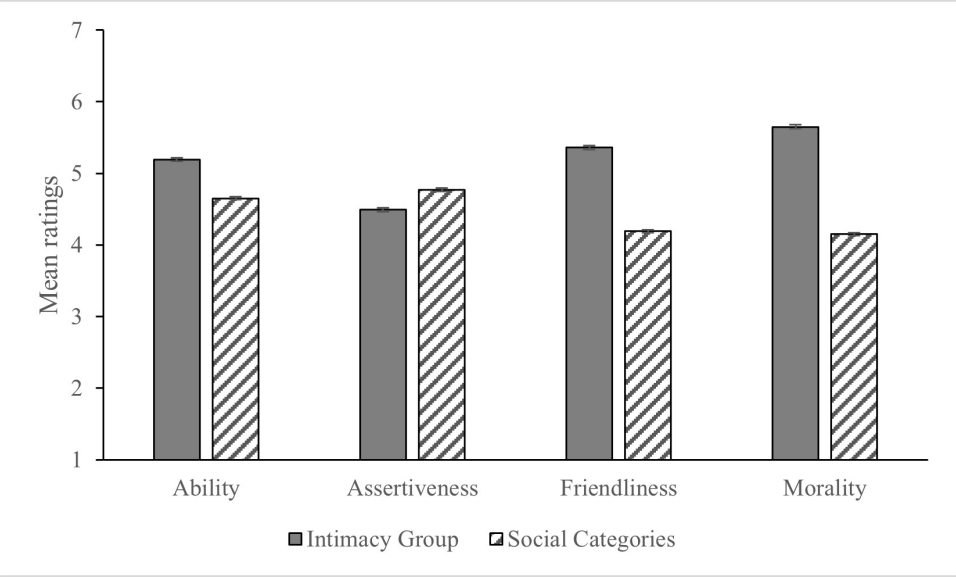

Figure 1a. Facet ratings as a function of the type of group in Belgium (Study 1a)

Note. Error bars represent standard errors.

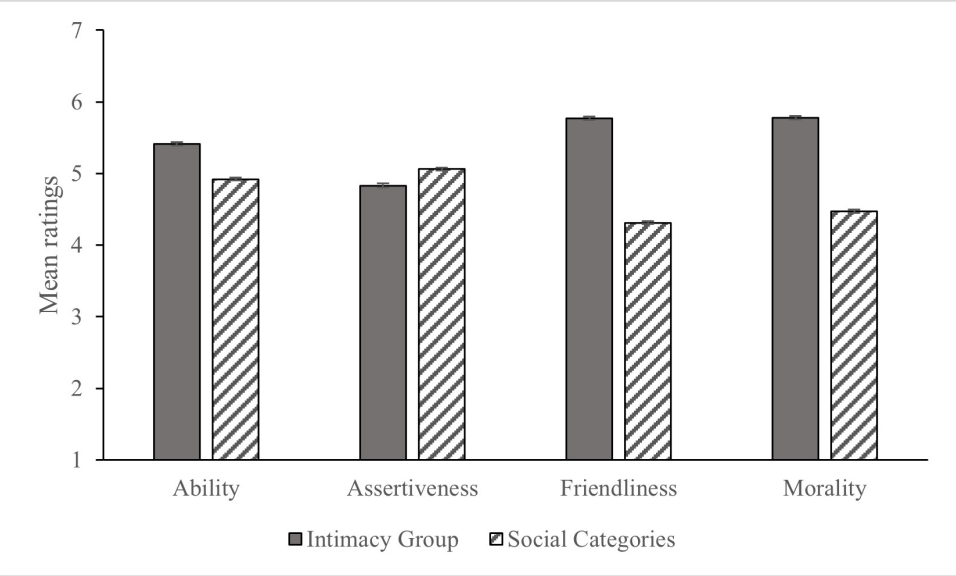

Figure 1b. Facet ratings as a function of the type of group in Germany (Study 1b)

Note. Error bars represent standard errors.

=.007, r^2 =.062) and morality (b = -1.30, $t(8)$ = -3.19, p =.013, r^2 =.05) for intimacy groups than for social categories.

Facet ratings within types of groups. In Study 1a (Belgium), intimacy groups were rated lower in assertiveness than in ability (b = -0.70, $t(18)$ = -6.20, p <.001, r^2 =.011), friendliness (b = -0.87, $t(18)$ = -7.73, p <.001, r^2 =.017), and morality (b = -1.16, $t(18)$ = -10.29, p <.001, r^2 =.030). Moreover, ability ratings of the intimacy group were significantly lower than morality ratings (b = 0.46, $t(18)$ = 4.09, p <.001, r^2 =.005), but not significantly different from friendliness ratings. Morality and friendliness ratings did not differ significantly.

Social categories in Study 1a (Belgium), were rated higher on assertiveness than on friendliness (b = -0.58, $t(17)$ = -5.25, p <.001, r^2 =.015) and on morality (b = -0.62, $t(17)$ = -5.64, p <.001, r^2 =.017). Ratings in ability were also higher than friendliness (b = -0.46, $t(17)$ = -4.14, p <.001, r^2 =.009), and morality ratings (b = -0.50, $t(17)$ = -4.54, p <.001, r^2 =.011).

Ability and assertiveness ratings as well as friendliness and morality ratings of social categories did not differ significantly.

Intimacy groups in Study 1b (Germany) were rated lower on assertiveness than on ability (b = -0.58, $t(20)$ = -6.66, p <.001, r^2 =.008), friendliness (b = -0.94, $t(20)$ = -10.83, p <.001, r^2 =.020), and morality (b = -0.95, $t(20)$ = -10.89, p <.001, r^2 =.020). Moreover, ability was rated lower than friendliness (b = -0.36, $t(20)$ = -4.16, p <.001, r^2 =.003) and morality (b = -0.37, $t(20)$ = -4.22, p <.001, r^2 =.003). Morality and friendliness ratings did not differ significantly for intimacy groups.

Social categories in Study 1b (Germany) were rated higher on assertiveness than on friendliness (b = -0.74, $t(17)$ = -8.88, p <.001, r^2 =.025) and morality (b = -0.58, $t(17)$ = -6.96, p <.001, r^2 =.015). Ability ratings were higher than friendliness (b = -0.61, $t(17)$ = -7.28, p <.001, r^2 =.017) and morality ratings (b = -0.45, $t(17)$ = -5.36, p <.001, r^2 =.009). Again, assertiveness and ability as well as morality and friendliness ratings did not differ significantly.[3]

Discussion

The first aim of Study 1 was to construct a measure of the two-dimension/four-facet model in group evaluation. Supporting our hypotheses, we found evidence for a better fit for models differentiating the evaluation on facets over models distinguishing only dimensions.

Moreover, the model proved acceptable both in the Belgian (French language) and in the German (German language) sample. With respect to measurement invariance, we found a satisfactory fit for configural invariance but not for metric invariance. This means that the two-dimension/four-facet model of social evaluation was applicable in both languages. At the same time, the specific traits show somewhat different loadings on the various facets in the two languages, thus preventing a direct comparison between the facet means obtained in Belgium and Germany.

The analyses also revealed that the facet scales showed acceptable reliabilities. With respect to the intercorrelations, assertiveness was most strongly correlated with ability, supporting our conceptualization of the Vertical dimension. Friendliness and morality were strongly correlated, supporting the conceptualization of the Horizontal dimension. Ability, however, was not only correlated with as-

3 Those results remained the same when controlling for other variables like the size of the group.

sertiveness but also with morality (and friendliness). This latter association between the Vertical (ability) and the Horizontal (morality) dimension is likely due to valence. As previous research showed (Abele, 2022), assertiveness is less positively evaluated than both ability and the Horizontal facets. We will come back to this issue in the General Discussion.

Regarding our second research question, we investigated the differences between two types of groups, intimacy groups, and social categories, as targets of evaluation. Confirming Lickel's (2000) taxonomy and our hypotheses, participants rated intimacy groups higher on entitativity, identification, similarity, and importance than social categories. More importantly, our findings showed that facet ratings differed depending on the type of group. Supporting our hypotheses, participants rated intimacy groups significantly higher on friendliness and morality than social categories. This pattern proved consistent across languages and is in line with the taxonomy in which intimacy groups are described to be "regulated by principles of generosity or communal sharing" (Lickel et al., 2000, p. 242).

Turning to differences within type of group, intimacy groups were higher on ability, friendliness, and morality than assertiveness, and higher on morality and friendliness than ability. Social categories received higher ratings on assertiveness and ability than on friendliness and morality. We note that, across types of groups, the ratings of assertiveness were most consistently different from those on other facets. This distinction of assertiveness in comparison to ability is of particular interest because this would not be reflected by a two-dimension approach. As such, this finding lends further credence to the relevance of a facet approach when investigating social evaluation. Interestingly, these differences in the ratings of the two types of groups replicated across both countries/languages.

To sum up, Studies 1a and 1b provided clear evidence that the two-dimension/four-facet model also emerges when it comes to the evaluation of groups. Our findings confirm that the distinction between the four facets helps to appraise groups on a more fine-grained level. We also found strong support for the usefulness of distinguishing between different types when evaluating groups.

One limitation of Study 1 is that participants provided ratings about intimacy groups and social categories but did not consider task groups. Because this is the third important type of group considered by Lickel et al. (2000, 2001), we examined this in Study 2.

Study 2

In Study 2, we wanted to turn to a different type of group, namely task groups with the aims to (1) replicate the relevance of our item list and structural model on task group, and (2) to observe how the group features and the facet ratings behave in task groups. We presented our participants with a definition of task groups and gave some examples before asking them to select and describe a task group in which they themselves had been or were still members. As was the case for the intimacy groups examined in Study 1, this approach ensured that participants thought about a real group and that there would be variability in the selected task groups. We again collected the data online in both Belgium (Study 2a) and Germany (Study 2b). As before, we excluded participants who were not German/French native speakers and those who considered a group comprising less than three members.

Study 2a: Belgium

In Belgium $N = 151$ students participated (110 female, 40 male, one diverse; age range from 18 to 47 years, $M = 20.8$, $SD = 4.07$). The majority held a secondary degree ($n = 119$, 78.8%; academic degree, $n = 32$, 21.2%).

Study 2b: Germany

In Germany, $N = 115$ participants took part (69 female, 45 male, one diverse; age range from 18 – 35 years, $M = 21.99$, $SD = 3.21$). The majority held an academic degree ($n = 82$, 71.3%; secondary degree, $n = 33$, 28.7%).

Method

Procedure and Measures

We asked participants to think of a task group of which they were or had been a member. We gave examples of task groups (e.g. project teams in business, student work groups) and asked them to describe the task group they had selected in a few sentences. Specifically, they had to indicate the rough size of the group, its duration, and whether they were still a member. Next, participants completed the same facet, entitativity, similarity, identification, and importance items as before.

Analyses

We conducted the same analyses as in Study 1 with principal component analyses (PCAs) as screening for the items, and CFAs for testing the fit of our proposed two-dimension/four-facet model. Then we analyzed the model with respect to task groups. Analysis codes and data for both Studies 2a and 2b are available via https://osf.io/3srq5/ on the Open Science Framework.

Results of Studies 2a and 2b

Scale construction and Confirmatory Factor Analyses

The PCAs resulted in the same item selection as in Study 1 (see supplementary material S2.5 - S2.6). Interestingly, the correlation between ability and morality was high (see Table 2). The CFAs showed acceptable levels of fit for the two-dimension/four-facet model with a CFI value of .93 in Study 2a (Belgium) and .88 in Study 2b (Germany). Fit indices can be found in Table 1. Similar to Study 1 and again confirming our hypotheses, the four-facet and two-dimension/four-facet models did not differ significantly from each other ($p > .05$) while always performing better than the two-dimension and the one-dimension models ($p < .05$).

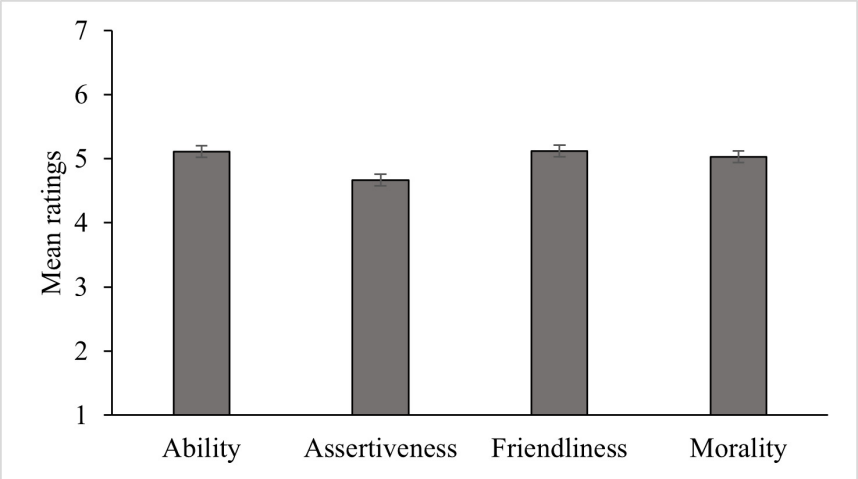

Figure 2a. Facet ratings of work groups in Belgium

Note. Error bars represent standard errors.

We again performed invariance analyses between both samples. We found a satisfactory fit in configural invariance ($\chi^2(330) = 3445.54$, $p < .001$, CFI =.91, RMSEA =.09, SRMR =.06) confirming that the data support the two-dimension/four-facet structure in both samples. The test of metric invariance statistically differed from the configural invariance test ($\Delta\chi^2$ (18) = 33.16, p =.016), signaling that we could not directly compare the data between samples. The reliabilities of the facet scales were good to excellent with alphas ranging from α =.81 to α =.92. In light of these results, we consolidated the facet scales obtained in Study 1.

Group features

We compared ratings of task groups' entitativity, similarity, identification, and importance. Mean values and standard deviations are displayed in Table 3. Although the mean values cannot be directly compared to those of Study 1, we note that they are always lower than the Study 1 mean ratings for intimacy groups, and higher than the Study 1 mean ratings for social categories. These differences are in line with the hypothesized position of task groups being located in between intimacy groups and social categories.

Facet evaluation

We conducted a mixed model with facet as a fixed factor varying within participants and participants and items as random intercepts. Figure 2 presents the results in both countries. For Study 2a (Belgium), we found no significant differences between the facet ratings. For Study 2b (Germany), friendliness ratings were lower than ability ratings (b = -0.84, $t(16)$ = -2.96, p =.009, r^2 =0.038) and there were no differences for the remaining facets.

Discussion

Study 2 investigated the facet approach for task groups. The analyses revealed a close link between ability and morality both in the factorial structure and in the evaluation. As much as this, the respective CFA in the two languages again showed acceptable levels of fit with a better fit for the two-dimension/four-facet model than for the two-dimension model, again stressing the relevance of a facet approach. With respect to the evaluation on the four facets, the results were slightly surprising. In Study 2a (Belgium), we found no differences in facet evaluation. In Study 2b (Germany), participants only evaluated task groups higher on ability than on friendliness.

The somewhat simpler factorial structure as well as the absence of more differentiation in the facet evaluation, could be due to the samples we relied on in Study 2 but also to the unique nature of task groups. Indeed, both samples comprised students who evaluated a self-chosen task group. Because students often change groups depending on the class, the year, and the type of assignment, the task groups evaluated might have been quite transitory. Additionally, the results of Study 2 suggest that the two-dimension/four-facet model may apply somewhat differently depending on the type of group. Whereas the model was well supported by the data in Study 1 in both Belgian and German samples and for both intimacy groups and social categories, the facet distinction may offer less on an increment compared to a two-dimension model when it comes to evaluating more competence-oriented groups such as task groups. Interestingly, we can note that the entitativity and further group feature ratings of the task groups of Study 2 descriptively fell between the ratings of the intimacy groups and social categories of Study 1. As such, this pattern supports our hypotheses and earlier findings on group typology (Lickel et al., 2000).

Study 3

Our third study pursued the same two aims as Study 1 and Study 2, namely (1) testing the two-dimension/four-facet model with respect to the evaluation of groups and establishing a reliable measurement scale, and (2) investigating how different types of groups may shape social evaluation. We also had several additional aims. First, we

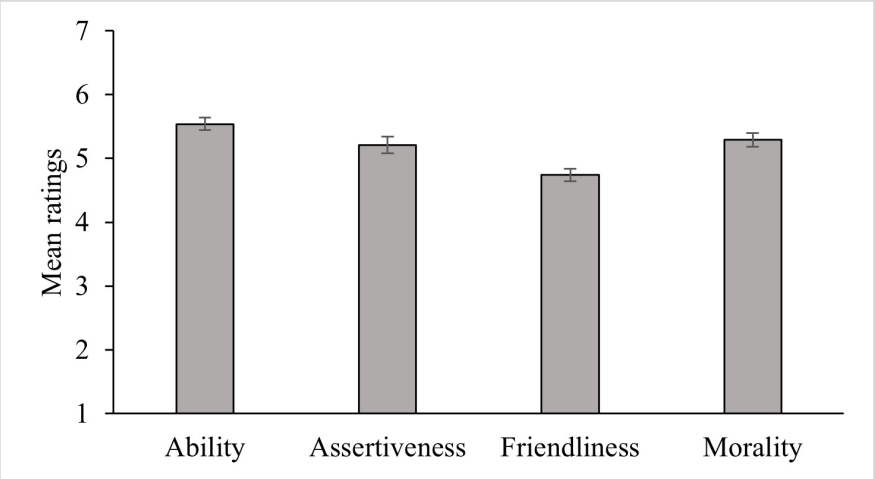

Figure 2b. Facet ratings of work groups in Germany

Note. Error bars represent standard errors.

wanted to extend the findings obtained in German and French to a third language and a third country. To this end, we conducted Study 3 in Britain with an English-speaking sample. Second, we wanted to have measures of all three types of groups within one sample. Hence, the participants of our third study rated an intimacy group, a task group, and a social category. Third, we wanted to make sure that all groups rated were "ingroups", that is, groups to which participants belonged. In the previous studies, intimacy groups and task groups were always ingroups whereas social categories were mostly outgroups. Because intimacy groups and task groups are naturally ingroups but social categories can be both ingroups and outgroups, we overcame the confound by showing participants only ingroups in Study 3. We presented participants with only one group of each type. We chose the "family" group (see Study 1) as an operationalization of intimacy groups, a self-chosen task group (see Study 2) as an operationalization of task groups, and we chose "British people" as the social category. The reason for choosing "British people" was that this social category was an ingroup for all our participants in Study 3. Indeed, our sample comprised only people living in Britain and self-defining as British. Furthermore, the national group was not present in Study 1 so we came up with a new social category in the present study to secure generalization. Along similar lines, "family" and self-selected "task group" are always ingroups. As a final change, we relied on a more comprehensive measure of entitativity in our Study 3.

In sum, Study 3 combines the designs of Studies 1 and 2 using more precise measures, more relevant target groups, and overcoming limitations of the previous studies. In addition, the British sample allowed us to establish our scale in another language and to test the robustness of our findings in another country. Our hypotheses were the same as in previous studies.

Method

Participants

The sample consisted of Prolific users who were born in the UK. They received one pound for participation. We excluded participants who failed the attention check, were not English native speakers, or did not show any variance in their responses. The final sample comprised N = 194 British people, half of which considered themselves as female (n = 97, 50% female; n = 95, 48.97 % male; n = 2, 1.03% other). The mean age was 41.1 years (SD = 13.66) with the youngest participant being 18 and the oldest being 73 years old. Most participants had a college or higher degree (n = 110, 56.70%), a smaller number had a secondary degree or vocational training (n = 84, 43.30%) and the majority was employed (n = 133, 68.56%).

Procedure and measures

The procedure was the same as in Studies 1 and 2. After providing consent, participants rated three groups (their family, a self-chosen task group, and British people) in random order.

Facet ratings. We relied on an English translation of the items used in Studies 1 and 2. Following the findings of principal component analyses (see supplementary material S2.7 – S2.9) we selected a set of 20 items. For the full set of items, see Appendix 1. All 7-point scales ranged from "1" (= not at all) to "7" (= very much).

Feature ratings. The three items referring to identification, similarity, and importance were the same as before. Because we aimed for a more comprehensive measure of group entitativity, we added three items to the one used in previous studies. That is, participants rated the extent to which they saw the group as unified, the similarity among the members in the group, how much the group looked like a group rather than just a bunch of individuals, and to what extent they considered the group to be a "real" group (see Callahan & Ledgerwood, 2016, for similar items). The reli-

ability analyses for the entitativity construct were good (intimacy group, α =.87; task group, α =.83; social category, α =.82). Regarding the task groups, participants also had to indicate the size of the group, its duration, and whether they were still members. Analysis codes and data for Study 3 are available via https://osf.io/3srq5/ on the Open Science Framework.

Results

Scale construction

We again analyzed the data using PCAs and CFAs. The exploratory factor analyses proved satisfactory (see Table S2.7 to S2.10). The obtained solution of the CFAs showed acceptable levels of fit (see Table 1). As before, the fit of the two-dimension model fared significantly better than that of the one-dimension model (p <.001). The four-facet and two-dimension/four-facet models did not differ significantly from each other (p >.05) while always performing better than the two-dimension model (p <.05). These findings again supported our first hypothesis on the applicability of the two-dimension/four-facet model to groups, as well as its superiority compared to a two-dimension model. The resulting facet scales showed good to very good reliabilities, ranging from α =.78 to α =.92.

Group features

We looked at the entitativity, similarity, identification, and importance ratings to test whether the three types of groups differed, in accordance with our hypotheses and previous research (Lickel et al., 2000). We conducted a mixed model with the type of group as a predictor and the respective ratings as the criterion. We treated participants as random intercepts. Table 3 show the ratings of groups' entitativity, similarity, identification, and importance. Confirming hypotheses, the intimacy group was highest on entitativity, followed by the task group, and then the social category, which had the lowest entitativity ratings. Identification, similarity, and importance ratings were always higher for the intimacy group than for the task group and the social category. The latter did not differ from each other.

Facet evaluation

Facet ratings. We computed a mixed model with type of group and facet, both varying within participants, and their interaction as fixed effects and both participants and items as random intercepts. Table 4 shows the respective means and standard deviations. The analysis revealed a significant main effect of facet, $F(3,16) = 4.65$, p =.016, with lower ratings of assertiveness than ability, $t(16) = -3.38$, p =.004, r^2 =.014, and morality , $t(16) = -3.07$, p =.007, r^2 =.011.

Types of group ratings. We also found a significant main effect of the type of group, $F(2, 11419) = 836.99$, p <.001. Follow-up analyses revealed higher average ratings for the intimacy group than the task group, $t(11425) = -7.44$, p <.001, r^2 =.003, and the social category, $t(11425) = -38.1$, p <.001, r^2 =.082. Moreover, we found higher ratings for the task group than the social category, $t(11425) = -30.6$, p <.001, r^2 =.055. These findings reveal a generally more positive evaluation of intimacy groups than of task groups, both being better evaluated than social categories.

Facet ratings between types of groups. The type of group by facet interaction also proved significant, $F(6, 11419) = 54.23$, p <.001. Comparing the type of groups, ability ratings were lower for the social category than for the intimacy group, $t(11419) = 18.80$, p <.001, r^2 =.021, and for the task group, $t(11419) = 18.46$, p <.001, r^2 =.021, but not different between intimacy group and task group.

Similarly, assertiveness ratings were lower for the social category than for the intimacy group, $t(11419) = 7.99$, p <.001, r^2 =.004, and for the task group, $t(11419) = 13.73$, p <.001, r^2 =.012, while being lower for the intimacy group than the task group, $t(11419) = 5.74$, p <.001, r^2 =.002.

Friendliness ratings were lower for the social than the intimacy group, $t(11419) = 21.01$, p <.001, r^2 =.027, and the task group, $t(11419) = 9.73$, p <.001, r^2 =.006, while being higher for the intimacy group than the task group, $t(11419) = -11.28$, p <.001, r^2 =.008.

Morality ratings were also lower for the social category than the intimacy group, $t(11419) = 29.39$, p <.001, r^2 =.051, and the task group, $t(11419) = 20.19$, p <.001, r^2 =.025 while being higher for the intimacy group than the task group, $t(11419) = -9.20$, p <.001, r^2 =.005.

These findings support our hypotheses. Intimacy groups are highest on the Horizontal dimension while the social category is lowest. Task groups are highest on the Vertical dimension, while the social category is lowest.

Facet ratings within types of groups. Participants rated the intimacy group lower on assertiveness than on ability, $t(19) = -5.23$, p <.001, r^2 =.013, friendliness, $t(19) = -5.51$, p <.001, r^2 =.015, and morality, $t(19) = -7.23$, p <.001, r^2 =.025. Participants did not evaluate the ability, friendliness, and morality of the intimacy group differently.

The task group also came out lower on assertiveness than on ability, $t(19) = -3.08$, p =.006, r^2 =.005, and higher on ability than on friendliness, $t(19) = -3.60$, p =.002, r^2 =.006. Ratings of assertiveness, friendliness, and morality did not differ significantly. The ratings of the social category did not differ significantly between the facets.

Discussion

In Study 3, we again aimed to validate a scale of items measuring the four facets of social evaluation with groups as targets but now in English besides French (Studies 1a and 2a) and German (Studies 1b and 2b). We asked each participant to evaluate one instance of each type of group. To address a potential limitation in previous studies, all targets were now ingroups, that is, groups to which participants belonged. Supporting our hypotheses, the results of Study 3 showed satisfactory fit indices for the two-dimension/four-facet model, and this model performed significantly better than a model only relying on two dimensions. However, consistent with Studies 1 and 2, the improvement for the task group was again less marked than for the other groups. The fact that the sample comprised non-student re-

Table 4. Ratings on facets depending on the type of group in study 3.

	Study 3			
	All groups	Intimacy group	Task group	Social category
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*
All facets		5.49 (1.11)	5.30 (1.11)	4.55 (0.92)
Ability	5.28 (1.07)	5.59 (0.96)	5.57 (1.09)	4.67 (0.89)
Assertiveness	4.83 (0.97)	4.87 (1.03)	5.15 (0.91)	4.48 (0.83)
Friendliness	5.10 (1.18)	5.62 (1.12)	5.08 (1.19)	4.61 (0.99)
Morality	5.24 (1.22)	5.86 (1.09)	5.41 (1.15)	4.43 (0.97)

spondents supports the assumption that the facets may apply to groups in slightly different ways depending on the type of groups.

The hypotheses on the different types of groups were also supported. The intimacy group was the highest on entitativity, identification, similarity, and importance, followed by the task group and the social category. Facet ratings differed between groups as predicted. Participants rates the intimacy group highest on the Horizontal dimension (both facets), and the task group highest on the Vertical dimension, particularly assertiveness. Within groups, participants rated the intimacy group lower on assertiveness than on the other facets. The task group was highest on ability and morality, and lower on friendliness and assertiveness. The social category did not show differences in facet ratings.

General Discussion

The present research focused on a novel approach to the social evaluation of groups. Specifically, we combined recent developments regarding the Big Two in social evaluation, particularly the facet approach (see Abele et al., 2021; Koch et al., 2021; Yzerbyt & Abele, 2022), with a systematic typology of groups that distinguishes between intimacy groups, task groups, and social categories (Lewin, 1948; Lickel et al., 2000, 2001). Evidently, previous research on the evaluation of groups has concentrated on the Big Two (Fiske et al., 2002; Judd et al., 2005; Yzerbyt et al., 2005), or has looked more precisely at competence, sociability, and morality (Brambilla et al., 2012; Ellemers, 2017; Ellemers et al., 2008; Leach et al., 2007), ability versus assertiveness (Carrier et al., 2014; Louvet et al., 2019; Yzerbyt et al., 2022) or agency, beliefs and communion (Koch et al., 2016). Previous research has also not systematically analyzed different types of groups. Here, we identified a gap in this area of research in that no previous work examined (a) the two-dimension/four-factor model with respect to groups and with respect to its superiority compared to a two-dimension model; (b) analyzed the four facets of the Big Two in combination while conducting (c) a systematic analysis of the different types of groups. The present research aimed to address these questions in a systematic way.

We conducted the present studies with two key goals in mind. First, to build a measure of the two-dimension/four-facet model that would hold also for the evaluation of groups (and not only for individuals) while being at least as suitable as the two-dimension model. Second, we wanted to show that different types of groups differ systematically in their evaluation on the facets. Next to these two main goals, our research agenda also aimed to check whether the findings would replicate in different languages. To this end, we conducted our studies in French (Belgium), German (Germany), and English (UK).

The model. In over two decades, research has emphasized two fundamental dimensions in social evaluation, now further broken down into four facets (Abele et al., 2008, 2016; Carrier et al., 2014; Yzerbyt et al., 2022). However, consensus on the meaning and measurement of these four facets is lacking. Our studies aimed to validate a set of traits covering these facets across different groups and languages. Our findings suggest the two-dimension/four-facet model is relevant in the social evaluation of groups, outperforming simpler models.

One might argue that we only established configural invariance, but not metric invariance between the models in different languages. We think that this is an inherent problem in comparing across languages, but not an issue specific to our research. Indeed, we showed that although we could not directly compare findings across languages, the findings with respect to facet evaluations in the different types of groups were remarkably similar.

The intercorrelations of the facet scales generally followed our theoretical reasoning with assertiveness being most strongly related to ability (Vertical dimension), and friendliness being most strongly related to morality (Horizontal dimension). However, in quite a few cases ability was also strongly related to friendliness and/or morality. In our view, this is not a shortcoming of the model, but remarkably underlines the utility of the facet approach. In the general two-dimensional model, assertiveness/ability belongs to the vertical, and friendliness/morality belong to the horizontal dimension. However, these relationships may change as a function of the specific target that is being examined. Indeed, the four facets are differently related to valence (Abele, 2022): Assertiveness is generally evaluated less positively than ability, friendliness, and morality. The intercorrelation of ability with the Horizontal facets may, in part, be due to this valence factor.

In addition, the factorial structure also varied with the target of evaluation. Across our studies, we found the best

fit indices for intimacy groups and social categories while the data for the task group structures seemed somewhat less impressive. That is, although the two-dimension/four-facet structure was always better than the two-dimension structure, the gain in χ^2 was systematically smaller for the task group than for the other types of groups (see Table S2.11). Indeed, we note that the distinction between friendliness and morality emerged less strongly in Study 2 when we asked students to evaluate task group. We observed the same pattern in Study 3, this time using a much more diverse sample. This supports the idea that the difference between the facet relations may depend more readily on the type of group rather than the sample. All in all, these observations emphasize the relevance of relying on a typology of groups in social evaluation studies.

The facets. The distinction of the facets per Big Two dimensions has already proven fruitful for the analysis of self-perception (Abele, 2022; Abele & Hauke, 2019), and of perception of others (Abele & Hauke, 2019; Louvet et al., 2019; Martinescu et al., 2022; Yzerbyt et al., 2022). As a case in point, research suggests that the perception of groups is often more related to morality than to friendliness (Brambilla et al., 2021; Leach et al., 2007). The present research extends these findings by examining the relevance of all four facets for the evaluation of three major group types distinguished by Lickel et al. (2000, 2001; see also Lewin, 1948). First of all, the group feature ratings of entitativity, similarity, importance, and identification showed that we successfully operationalized our three types of groups. Second, supporting our hypotheses on differences in ratings between group types, we found that participants rated intimacy groups higher than task groups than social categories on the Horizontal dimension. Participants also rated task groups higher on the Vertical dimension than intimacy groups, and the latter higher than social categories. Most importantly, we found that these general results are further refined when we capitalize on the facet ratings. Specifically, regarding the Vertical dimension, participants rated intimacy groups (Study 1, Study 3) and task groups (Study 3) higher on ability than on assertiveness. The absence of a difference in ability and the presence of a difference in assertiveness between the intimacy group and the task group in Study 3 again emphasizes the relevance of the facet approach. These findings show that the facet distinction is helpful in describing different types of groups. This is particularly the case for ability and assertiveness which led to different ratings in all three types of groups.

Taken together, both the analysis and comparison of the measurement models as well as the analysis of differences between and within the three types of groups suggest that the distinction into four facets helps to study in more detail the evaluation of groups in general, and of various group types in particular.

Theoretical Implications

The present research sends an important message with respect to the evaluation of groups. Indeed, we found convergent evidence from three studies that the same traits used to tap the Big Two and their associated facets can be used to evaluate different types of groups. This suggests that the vertical and horizontal dimensions and the facets of assertiveness, ability, friendliness, and morality constitute meaningful tools to appraise the way social perceivers experience the diverse groups that comprise their environment. Whereas the horizontal dimension is most important for the evaluation of intimacy groups, the vertical dimension appears quite relevant for the evaluation of task groups and social categories. Whereas the morality and the friendliness facets of the horizontal dimension are of similar importance in the evaluation of groups, the ability and assertiveness facets of the vertical dimension would seem to serve different functions. Specifically, judgments of ability tend to be associated with positive group relations while judgments of assertiveness appear to be associated with more psychological distance (social categories in Study 1) or with task orientation (Study 3). It would be interesting to more systematically link the structural (entitativity) and relational (liking, closeness, etc.) characteristics of groups to their evaluation on the facets.

Limitations and Future Directions

To be sure, the present work also comes with some limitations. A first question concerns the fact that we tested the two-dimension/four-facet model in a limited number of countries, using three languages. Although many other studies have been conducted in one language only, it remains of course a valuable topic for future research to examine the cross-cultural relevance of the model. We relied on three Western European samples so it is important to see if the same model holds in Asian, American, or African cultures. In light on research on person evaluation conducted in a large set of different countries (Abele et al., 2016), we surmise that the facets are also psychologically meaningful for group evaluation, although empirical evidence is of course crucial in this respect.

A second issue is that our participants evaluated a limited number of groups from each type. We had two exemplars of intimacy groups (family and friends), and it is an open question if there are more exemplars of intimacy groups. We had self-selected task groups, hence, there was variation in this duration of type of group. Finally, we had nine exemplars of social categories (eight in Study 1, preselected from Koch et al., 2016; one in Study 3). Of course, one can think of many more social categories, and future research could do a more fine-grained analysis of the evaluation of social categories.

A third issue arising from our data and a clear avenue for further research has to do with a systematic variation of group type by ingroup–outgroup membership. It may well be that superimposing the ingroup–outgroup distinction above the group type distinction, i.e. intimacy groups, task groups, and social categories to which a person does or does not belong, helps to clarify the association of the facets with group evaluation. One could, for instance, test if the lower importance of friendliness than morality in evaluating groups (Brambilla et al., 2012; Ellemers, 2017; Ellemers et al., 2008; Leach et al., 2007) is confined to outgroups or groups to which the evaluator does not (yet) belong, while

friendliness becomes more important for ingroups, particularly task groups of which a person is a member. As another example, assertiveness may be more critical than ability for outgroups whereas this difference is less relevant or even reverses when it comes to ingroups (Castano et al., 2003; Yzerbyt et al., 2000, 2022). Conversely, assertiveness may become more important in ingroups than outgroups, when group members' outcomes are dependent on the group's ability and confidence (Abele & Brack, 2013).

A fourth issue that could be the focus of future work concerns group features and their association with the facets. In the present endeavor, we decided to shed light specifically on entitativity, importance, similarity, and identification. This list could be extended with, for instance, such characteristics as power, influence, and status of groups (Yzerbyt et al., 2022), or, alternatively, cohesion, and solidarity of groups could all be the focus of future studies. An important future step can be to manipulate some selected group characteristics, say entitativity, and examine the way this influences the evaluation of the groups on the facets of the Big Two. Indeed, in light of the role of the two dimensions and the facets in people's experience of groups, a better knowledge of the impact of certain group features on the perception of the facets would be of paramount importance. These questions may become critical in terms of group impression management insofar as group members may want to monitor certain features to ensure that social perceivers appraise their group in specific ways.

Another avenue for future work concerns the relation between the type of group and the factorial structure of the facets. Our results suggest that the facets may be easy to distinguish when evaluating intimacy groups and social categories but that they tend to overlap somewhat when evaluating task groups. A closer examination of this phenomenon may prove valuable and further our understanding of the inter-facet dynamics in social evaluation.

A final topic for further research could be the interrelationship of the facets with global valence. This issue seems to be of special relevance for the assertiveness facet since this facet seems to come across as less positive than the other three facets. Interestingly enough, this facet is strongly associated with a series of positive outcomes for the self like status, economic success, self-efficacy, or self-esteem (Abele, 2022; Abele & Hauke, 2019). More work is thus needed to delineate the relationship between assertiveness and valence.

Practical Implications

The lessons learned from the present efforts also entail several practical implications. Clearly, many studies are starting to look at social evaluation of groups using the framework of the Big Two and the four facets (e.g., Koch et al., 2021; Nicolas, Bai, et al., 2022; Yzerbyt et al., 2022). Researchers often rely on ad hoc lists of traits to measure dimensions, which can compromise comparability. The studies presented here offer a solid measurement tool for evaluating facets across diverse groups. Although minor adaptations may be needed for specific languages or specific targets, this should not prevent from using the set as it stands. The situation is comparable to what researchers encountered when relying on the semantic differential (Osgood, 1968). There too, there was some possibility that one of the scales would prove less than ideal when applied to a specific judgment target. However, this did not jeopardize the relevance of the tool. We would argue that the same holds here and would thus recommend using the same set of traits stemming from the present studies for all groups in all languages, only adapting things marginally to accommodate for exceptions. Keeping the instrument as standard as possible should not only allow researchers to gain a fine-grained understanding of the way any given group stands on the four facets and on the two dimensions, but they should also allow comparison across different groups.

Finally, a practical implication of the present research is the fact that the type of group matters. Quite a few inconsistencies in the research field of the Big Two in group evaluation could be reconciled if one would consider the type of group examined. Indeed, given the fact that social categories are far from being the sole or even the most important type of group that people may face in their daily lives, one needs to be careful about what is most meaningful for the respondent in the specific study at hand.

Conclusions

Our research presented a model that combined recent developments in the field of social evaluation (Abele et al., 2021; Koch et al., 2021), i.e. the Big Two and their facets, with a typology of groups that distinguishes three main types of intimacy groups, task groups, and social categories (Lickel et al., 2000, 2001). Besides constructing a measure of the Big Two and their facets and validating it in three different languages (English, French, German) the present research showed that our model (a) revealed deeper insights into the dynamics of group evaluation; (b) helps to reconcile inconsistent findings in the literature and, hence, to extend existing models of group evaluation; and (c) opens various avenues for further research on group evaluation, for instance, with respect to group type and ingroup – outgroup differentiation, or with respect for different group features, like, for instance, entitativity, and how these are related to the facets of the Big Two.

Contributions

Contributed to conception and design: JB, JS, VY, AA
Contributed to the acquisition of data: JB, JS
Contributed to analysis and interpretation of data: JB, JS, VY, AA
Drafted and/or revised the article: JB, JS, VY, AA
Approved the submitted version for publication: JB, JS, VY, AA

Acknowledgments

We thank Laura Gabel for her help in recruiting participants and collecting data for study 2b.

Funding Information

Vincent Yzerbyt was supported by a Templeton World Charity Foundation grant (2022-30553)

The collection of the Belgian data was supported by a Fonds De La Recherche Scientifique - FNRS grant (PDR T.0253.19) to Vincent Yzerbyt.

The collection of the German data was supported by a Robert Bosch grant (B-107433) to Andrea Abele.

Competing Interests

The authors have no conflict of interest to declare.

Data Accessibility Statement

The data and code are available on the Open Science Framework at https://osf.io/3srq5/.

Submitted: May 14, 2024 PDT, Accepted: September 03, 2024 PDT

References

Abele, A. E. (2022). Evaluation of the Self on the Big Two and their Facets: Exploring the Model and its Nomological Network. *International Review of Social Psychology*, *35*(1), Article 14. https://doi.org/10.5334/irsp.688

Abele, A. E., & Brack, S. (2013). Preference for Other Persons' Traits Is Dependent on the Kind of Social Relationship. *Social Psychology*, *44*(2), 84–94. https://doi.org/10.1027/1864-9335/a000138

Abele, A. E., Cuddy, A. J. C., Judd, C. M., & Yzerbyt, V. Y. (2008). Fundamental dimensions of social judgment. *European Journal of Social Psychology*, *38*(7), 1063–1065. https://doi.org/10.1002/ejsp.574

Abele, A. E., Ellemers, N., Fiske, S. T., Koch, A., & Yzerbyt, V. (2021). 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. (2019). Comparing the facets of the big two in global evaluation of self versus other people. *European Journal of Social Psychology*, *50*(5), 969–982. https://doi.org/10.1002/ejsp.2639

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*, 1810. 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*(5), 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. In *Advances in Experimental Social Psychology* (Vol. 50, pp. 195–255). Elsevier. https://doi.org/10.1016/B978-0-12-800284-1.00004-7

Ægisdóttir, S., Gerstein, L. H., & Çinarbaş, D. C. (2008). Methodological Issues in Cross-Cultural Counseling Research. *The Counseling Psychologist*, *36*(2), 188–219. https://doi.org/10.1177/0011000007305384

Augoustinos, M., & Walker, I. (2001). *Social cognition: An integrated introduction*. SAGE.

Brambilla, M., Sacchi, S., Rusconi, P., Cherubini, P., & Yzerbyt, V. Y. (2012). You want to give a good impression? Be honest! Moral traits dominate group impression formation. *The British Journal of Social Psychology*, *51*(1), 149–166. https://doi.org/10.1111/j.2044-8309.2010.02011.x

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). Elsevier. https://doi.org/10.1016/bs.aesp.2021.03.001

Brehm, S. S., Kassin, S. M., & Fein, S. (1999). *Social psychology* (4th ed.). Houghton Mifflin.

Brown, R. (Ed.). (1998). *Group processes: Dynamics within and between groups*. Wiley Blackwell. https://doi.org/10.1002/9781118719244

Burkley, E., Durante, F., Fiske, S. T., Burkley, M., & Andrade, A. (2017). Structure and content of Native American stereotypic subgroups: Not just (ig)noble. *Cultural Diversity & Ethnic Minority Psychology*, *23*(2), 209–219. https://doi.org/10.1037/cdp0000100

Callahan, S. P., & Ledgerwood, A. (2016). On the psychological function of flags and logos: Group identity symbols increase perceived entitativity. *Journal of Personality and Social Psychology*, *110*(4), 528–550. https://doi.org/10.1037/pspi0000047

Campbell, D. T. (1958). Common fate, similarity, and other indices of the status of aggregates of persons as social entities. *Behavioral Science*, *3*(1), 14–25. https://doi.org/10.1002/bs.3830030103

Carnes, N. C., Lickel, B., & Janoff-Bulman, R. (2015). Shared perceptions: Morality is embedded in social contexts. *Personality & Social Psychology Bulletin*, *41*(3), 351–362. https://doi.org/10.1177/0146167214566187

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

Castano, E., Yzerbyt, V., & Bourguignon, D. (2003). We are one and I like it: The impact of ingroup entitativity on ingroup identification. *European Journal of Social Psychology*, *33*(6), 735–754. https://doi.org/10.1002/ejsp.175

Cooley, C. H. (1909). *Social organization: A study of the larger mind*. Charles Scribner's Sons. https://doi.org/10.1037/14788-000

Crawford, M. T., & Salaman, L. (2012). Entitativity, identity, and the fulfillment of psychological needs. *Journal of Experimental Social Psychology*, *48*(3), 726–730. https://doi.org/10.1016/j.jesp.2011.12.015

Cuddy, A. J. C., Fiske, S. T., Kwan, V. S. Y., Glick, P., Demoulin, S., Leyens, J.-P., Bond, M. H., Croizet, J.-C., Ellemers, N., Sleebos, E., Htun, T. T., Kim, H.-J., Maio, G., Perry, J., Petkova, K., Todorov, V., Rodríguez-Bailón, R., Morales, E., Moya, M., … Ziegler, R. (2009). Stereotype content model across cultures: Towards universal similarities and some differences. *British Journal of Social Psychology*, *48*(1), 1–33. https://doi.org/10.1348/014466608X314935

Denson, T. F., Lickel, B., Curtis, M., Stenstrom, D. M., & Ames, D. R. (2006). The Roles of Entitativity and Essentiality in Judgments of Collective Responsibility. *Group Processes & Intergroup Relations*, *9*(1), 43–61. https://doi.org/10.1177/1368430206059857

Ellemers, N. (2017). *Morality and the Regulation of Social Behavior*. Routledge. https://doi.org/10.4324/9781315661322

Ellemers, N., Pagliaro, S., Barreto, M., & Leach, C. W. (2008). Is it better to be moral than smart? The effects of morality and competence norms on the decision to work at group status improvement. *Journal of Personality and Social Psychology*, *95*(6), 1397–1410. https://doi.org/10.1037/a0012628

Fiske, S. T. (2015). Intergroup Biases: A Focus on Stereotype Content. *Current Opinion in Behavioral Sciences*, *3*, 45–50. https://doi.org/10.1016/j.cobeha.2015.01.010

Fiske, S. T. (2018). Stereotype Content: Warmth and Competence Endure. *Current Directions in Psychological Science*, *27*(2), 67–73. https://doi.org/10.1177/0963721417738825

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, *11*(2), 77–83. https://doi.org/10.1016/j.tics.2006.11.005

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*(6), 878–902. https://doi.org/10.1037/0022-3514.82.6.878

Hamilton, D. L., Sherman, S. J., & Castelli, L. (2002). A Group By Any Other Name—The Role of Entitativity in Group Perception. *European Review of Social Psychology*, *12*(1), 139–166. https://doi.org/10.1080/14792772143000049

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Imhoff, R., Koch, A., & Flade, F. (2018). (Pre)occupations: A data-driven model of jobs and its consequences for categorization and evaluation. *Journal of Experimental Social Psychology*, *77*, 76–88. https://doi.org/10.1016/j.jesp.2018.04.001

Judd, C. M., James-Hawkins, L., Yzerbyt, V., & Kashima, Y. (2005). Fundamental dimensions of social judgment: Understanding the relations between judgments of competence and warmth. *Journal of Personality and Social Psychology*, *89*(6), 899–913. https://doi.org/10.1037/0022-3514.89.6.899

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., Imhoff, R., Unkelbach, C., Nicolas, G., Fiske, S., Terache, J., Carrier, A., & Yzerbyt, V. (2020). Groups' warmth is a personal matter: Understanding consensus on stereotype dimensions reconciles adversarial models of social evaluation. *Journal of Experimental Social Psychology*, *89*. https://doi.org/10.1016/j.jesp.2020.103995

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* (Vol. 63, pp. 1–68). Elsevier. https://doi.org/10.1016/bs.aesp.2020.11.001

La Macchia, S. T., Louis, W. R., Hornsey, M. J., & Leonardelli, G. J. (2016). In Small We Trust: Lay Theories About Small and Large Groups. *Personality and Social Psychology Bulletin*, *42*(10), 1321–1334. https://doi.org/10.1177/0146167216657360

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

Lewin, K. (1948). *Resolving social conflicts*. Harper.

Lickel, B., Hamilton, D. L., & Sherman, S. J. (2001). Elements of a Lay Theory of Groups: Types of Groups, Relational Styles, and the Perception of Group Entitativity. *Personality and Social Psychology Review*, *5*(2), 129–140. https://doi.org/10.1207/s15327957pspr0502_4

Lickel, B., Hamilton, D. L., Wieczorkowska, G., Lewis, A., Sherman, S. J., & Uhles, A. N. (2000). Varieties of groups and the perception of group entitativity. *Journal of Personality and Social Psychology*, *78*(2), 223–246. https://doi.org/10.1037//0022-3514.78.2.223

Lickel, B., Rutchick, A. M., Hamilton, D. L., & Sherman, S. J. (2006). Intuitive theories of group types and relational principles. *Journal of Experimental Social Psychology*, *42*(1), 28–39. https://doi.org/10.1016/j.jesp.2005.01.007

Louvet, E., Cambon, L., Milhabet, I., & Rohmer, O. (2019). The relationship between social status and the components of agency. *The Journal of Social Psychology*, *159*(1), 30–45. https://doi.org/10.1080/00224545.2018.1441795

Martinescu, E., Peters, K., & Beersma, B. (2022). What Do We Talk About When We Talk About Others? Evidence for the Primacy of the Horizontal Dimension of Social Evaluation in Workplace Gossip. *International Review of Social Psychology*, *35*(1), 13. https://doi.org/10.5334/irsp.687

Miraucourt, D., Caruana, S., & Mollaret, P. (2022). How Do People Evaluate Themselves in Terms of Assertiveness and Ability After Having Failed or Succeeded: The (Economic) Consequences Matter! *International Review of Social Psychology*, *35*(1), Article 17. https://doi.org/10.5334/irsp.692

Nicolas, G., Bai, X., & Fiske, S. T. (2022). A spontaneous stereotype content model: Taxonomy, properties, and prediction. *Journal of Personality and Social Psychology*. Advance online publication. https://doi.org/10.1037/pspa0000312

Nicolas, G., Fiske, S., Terache, J., Carrier, A., Yzerbyt, V. Y., Koch, A., Imhoff, R., & Unkelbach, C. (2022). Relational versus structural goals prioritize different social information. *Journal of Personality and Social Psychology*. Advance online publication. https://doi.org/10.1037/pspi0000366

Osgood, C. E. (1968). *Interpersonal verbs and interpersonal behavior*.

Plötner, M., Over, H., Carpenter, M., & Tomasello, M. (2016). What Is a Group? Young Children's Perceptions of Different Types of Groups and Group Entitativity. *PloS One*, *11*(3), e0152001. https://doi.org/10.1371/journal.pone.0152001

Sherif, M., & Sherif, C. W. (1956). *An outline of social psychology*. Harper & Bros.

Sherman, S. J., Castelli, L., & Hamilton, D. L. (2002). The spontaneous use of a group typology as an organizing principle in memory. *Journal of Personality and Social Psychology*, *82*(3), 328–342. https://doi.org/10.1037//0022-3514.82.3.328

Spencer-Rodgers, J., Hamilton, D. L., & Sherman, S. J. (2007). The central role of entitativity in stereotypes of social categories and task groups. *Journal of Personality and Social Psychology*, *92*(3), 369–388. https://doi.org/10.1037/0022-3514.92.3.369

Wojciszke, B., & Abele, A. E. (2018). Agency and communion in social cognition. In A. E. Abele & B. Wojciszke (Eds.), *Agency and Communion in Social Psychology*. Routledge. https://doi.org/10.4324/9780203703663-3

Yzerbyt, V. (2016). Intergroup stereotyping. *Current Opinion in Psychology*, *11*, 90–95. https://doi.org/10.1016/j.copsyc.2016.06.009

Yzerbyt, V., & Abele, A. (Eds.). (2022). The facets of social evaluation: The diversity of aspects underlying the Big Two of social perception [Special issue]. *International Review of Social Psychology*, *35*(1).

Yzerbyt, V., Barbedor, J., Carrier, A., & Rohmer, O. (2022). The Facets of Social Hierarchy: How Judges' Legitimacy Beliefs and Relative Status Shape Their Evaluation of Assertiveness and Ability. *International Review of Social Psychology*, *35*(1), 18. https://doi.org/10.5334/irsp.695

Yzerbyt, V., & Cambon, L. (2017). The Dynamics of Compensation: When Ingroup Favoritism Paves the Way for Outgroup Praise. *Personality and Social Psychology Bulletin*, *43*(5), 587–600. https://doi.org/10.1177/0146167216689066

Yzerbyt, V., Castano, E., Leyens, J.-P., & Paladino, M.-P. (2000). The Primacy of the Ingroup: The Interplay of Entitativity and Identification. *European Review of Social Psychology*, *11*(1), 257–295. https://doi.org/10.1080/14792772043000059

Yzerbyt, V., Corneille, O., & Estrada, C. (2001). The Interplay of Subjective Essentialism and Entitativity in the Formation of Stereotypes. *Personality and Social Psychology Review*, *5*(2), 141–155. https://doi.org/10.1207/S15327957PSPR0502_5

Yzerbyt, V., & Demoulin, S. (2010). Intergroup Relations. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of Social Psychology*. John Wiley & Sons, Inc. https://doi.org/10.1002/9780470561119.socpsy002028

Yzerbyt, V., Judd, C. M., & Corneille, O. (2004). *The psychology of group perception: Perceived variability, entitativity, and essentialism*. Psychology Press. https://doi.org/10.4324/9780203644973

Yzerbyt, V., Provost, V., & Corneille, O. (2005). Not Competent but Warm… Really? Compensatory Stereotypes in the French-speaking World. *Group Processes & Intergroup Relations*, *8*(3), 291–308. https://doi.org/10.1177/1368430205053944

Yzerbyt, V., Rocher, S., & Schadron, G. (1997). Stereotypes as explanations: A subjective essentialistic view of group perception. In R. Spears, P. J. Oakes, N. Ellemers, & S. A. Haslam (Eds.), *The social psychology of stereotyping and group life* (pp. 20–50). Blackwell Publishing.

Appendix

A1 *List of selected items*

(a) In French:

- Ability: *capable, compétent, doué, ingénieux,* and *intelligent*
- Assertiveness: *ambitieux, assertif, décidé, déterminé,* and *sûr de soi*
- Friendliness: *attentionné, chaleureux, cordial, sociable,* and *sympathique*
- Morality: *digne de confiance, fiable, honnête, sincère,* and *moral*

(b) In German:

- Ability: *clever, intelligent, kompetent, leistungsfähig,* and *schlau*
- Assertiveness: *durchsetzungsfähig, entschlossen, erfolgsorientiert, selbstsicher,* and *zielstrebig*
- Friendliness: *freundlich, fürsorglich, herzlich, liebevoll,* and *sympathisch*
- Morality: *ehrlich, fair, moralisch, vertrauenswürdig,* and *zuverlässig*

(c) In English:

- Ability: *capable, competent, intelligent, skillful,* and *smart*
- Assertiveness: *ambitious, assertive, decisive, determined,* and *self-assured*
- Friendliness: *affectionate, caring, friendly, sociable,* and *warm*
- Morality: *fair, honest, moral, reliable,* and *trustworthy*

Supplementary Materials

Supplemental Material

Download: https://collabra.scholasticahq.com/article/124471-a-novel-approach-to-the-evaluation-of-groups-type-of-group-and-facet-of-evaluation-matter/attachment/248565.docx?auth_token=zKTwWJ0U6YkDehtkkp6Q

Peer Review Communication

Download: https://collabra.scholasticahq.com/article/124471-a-novel-approach-to-the-evaluation-of-groups-type-of-group-and-facet-of-evaluation-matter/attachment/248566.docx?auth_token=zKTwWJ0U6YkDehtkkp6Q