

Research



Cite this article: Spadaro G, Jin S, Balliet D. 2022 Gender differences in cooperation across 20 societies: a meta-analysis. *Phil. Trans. R. Soc. B* **378**: 20210438. <https://doi.org/10.1098/rstb.2021.0438>

Received: 14 March 2022
Accepted: 17 June 2022

One contribution of 17 to a theme issue ‘Cooperation among women: evolutionary and cross-cultural perspectives’.

Subject Areas:
behaviour

Keywords:
cooperation, gender, meta-analysis, social dilemmas

Author for correspondence:
Giuliana Spadaro
e-mail: g.spadaro@vu.nl

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.6249964>.

Gender differences in cooperation across 20 societies: a meta-analysis

Giuliana Spadaro, Shuxian Jin and Daniel Balliet

Department of Experimental and Applied Psychology, Vrije Universiteit Amsterdam, Van der Boerhorststraat 7, 1081, Amsterdam, The Netherlands

ID GS, 0000-0002-5706-5898; SJ, 0000-0003-2209-4311; DB, 0000-0001-5365-9675

Past research hypothesized that men and women differ in their tendency to cooperate with strangers in situations that involve a conflict of interests. However, recent empirical research has provided converging evidence that men and women cooperate to a similar extent, and that differences in cooperation can emerge in response to specific situational and societal contexts. Here we analyse six decades of empirical research on human cooperation using social dilemmas (1961–2017, $k = 126$) conducted across 20 industrialized societies, testing pre-registered hypotheses derived from evolutionary theory and social role theory. Overall, our findings revealed little-to-no evidence for an association between gender and cooperation using different meta-analytic approaches. We did not find within-study differences in cooperation between men and women ($d = 0.011$, 95% CI $[-0.038, 0.060]$). However, cooperation was slightly higher across studies with predominantly female samples ($k = 972$). In addition, contrary to our predictions, gender differences in cooperation did not emerge in response to the degree of conflicting interests in the situation, and societal levels of gender equality and economic development. We discuss the implications of these findings for our understanding of gender differences in cooperation.

This article is part of the theme issue ‘Cooperation among women: evolutionary and cross-cultural perspectives’.

1. Introduction

Modern societies have pervasive inequalities in economic and social outcomes between men and women. Such disparities result in less opportunities for women to hold positions of economic and political power [1], and have often been explained by the assumption that men and women differ in their cooperative attitudes. Widely shared beliefs about gender roles ascribe women with more other-oriented, communal, and caring attributes, which convey lower social status than the competitive, agentic and self-enhancing attributes ascribed to men [2]. These gender-typical attributes might predict behaviour in situations of conflict between self-interest and collective welfare, and might contribute to reinforce and legitimize existing inequalities [3]. However, existing empirical evidence supporting gender differences in cooperation is mixed, including studies showing that men cooperate more than women [4–6], and also no gender differences at all [7,8], supporting the more general assumption that men and women respond differently to specific features of the context of the interaction [9]. Thus, it becomes key to understand not only whether, but also in which context such gender differences in cooperation may emerge.

To answer these questions, in this work we meta-analysed data from six decades of empirical studies conducted between 1961 and 2017 examining the relationship between participants’ gender and cooperation in highly standardized social dilemma paradigms where individuals make costly decisions to cooperate to increase collective welfare under specific interdependent payoff structures (described in detail in the methods section). We aimed to

(i) extend research on the contextual factors underlying gender differences in cooperation, (ii) examine the role of the broader cultural and institutional context, and (iii) replicate previous meta-analytic work on gender differences in cooperation.

Over the last decades, research investigating gender differences in cooperation relied on prominent theoretical frameworks to understand whether men and women respond differently to social dilemmas and whether any differences emerge in response to specific contextual factors: social role theory and evolutionary theory. According to social role theory, gender differences in social behaviour originate from the interplay between physical specialization between the sexes (e.g. reproductive abilities) and social influences [10]. Individuals are exposed to gender-specific family and employment role configurations in which men are primarily involved in agentic tasks and women in communal activities [2]. Such experiences are then internalized and reinforced through socialization processes, giving rise to shared normative expectations for men that conform to the society prevailing division of labour [11]. According to this perspective, women are hypothesized to be more cooperative than men in situations with a communal focus, because of the characteristics of gender-related stereotypes (i.e. higher communality and care for others' welfare) [12], while men tend to display prosocial behaviour in situations with an agentic focus (e.g. that provide opportunity to secure status) [2]. Furthermore, such gender differences based on social roles might be reduced over time while societies transition toward more egalitarian role configurations. Although women's greater communal orientation has been mainly theorized to apply to close relationships [13], such hypothesis can be extended to other-oriented behaviours toward strangers as well [12].

According to evolutionary accounts, gender differences in social behaviours emerged from the adaptive strategies that men and women respectively evolved in response to specific challenges and selective pressures faced in their ancestral environment [14]. For example, the ability to establish and rely on strong, large, coalitions of same-sex individuals could have been an adaptive strategy to succeed in sex-differentiated ancestral challenges such as hunting, trading or warfare [15,16]. Thus, interactions with same-sex partners [17], prolonged over time [18], and involving a large number of individuals [15] could prompt greater cooperation for men as compared to women. Predictions derived from these two theoretical frameworks have been tested in a meta-analysis that found greater cooperation from men in same-sex and repeated interactions, but greater cooperation from women in larger groups and more recent studies [12]. However, these theories make predictions about additional situational and societal dimensions that might moderate the relation between gender and cooperation that have not yet been tested.

The degree of conflict of interests characterizing the situation is a feature that can elicit differential responses in cooperation from men and women. Situations involving high conflict of interests expose individuals to greater vulnerability for cooperative behaviours to be exploited, as relatively higher gains can be achieved from defecting over cooperating. In fact, recent research has consistently observed lower cooperation in interactions involving non-corresponding interests (see [19] for a recent meta-analysis) and trust to be especially relevant in these situations [20]. However, differential cooperative responses in situations that expose to greater vulnerability could be expected

in men and women in light of both evolutionary selection pressures and gender identity processes. Specifically, responses to situations of vulnerability and stress have been hypothesized to have evolved resembling a 'tend and befriend' pattern from women, characterized by a tendency to enact conciliatory and affiliative behaviours [21]. As such, cooperative responses toward non-kin, compared to men-specific aggressive or avoidant coping reactions (i.e. 'fight or flight' patterns, [22]), could have been adaptive to maximize the survival of the offspring [21,23]. Thus, women may cooperate more than men in situations that involve a stronger conflict of interests, which pose a higher threat of exploitation.

The same hypothesis can also be posed from social role theory. In social dilemma studies, both men and women expect women to behave more cooperatively than men (e.g. [24]), and women are aware of these expectations (e.g. [25,26]). Notably, gender identity can be activated by salient situational cues and elicit subsequent gender-typical behaviour [27]. In fact, more cooperation is observed from women, compared to men, in groups that are heterogeneous with respect to gender [28]. Given the pervasive stereotype of women being more cooperative and men more competitive, exposure to competitive payoff structures could activate gender stereotypes, leading women to self-regulate behaviours according to the stereotype [3]. Evidence from economic games supports the hypothesis that gender differences might primarily emerged in situations involving greater conflict of interests [29], showing that women cooperated significantly more than men only in games that involved a stronger potential for exploitation (Prisoner's Dilemmas versus ultimatum games [30]) [31].

In this work, we tested this hypothesis operationalizing the degree of conflicting interests in Prisoner's Dilemma and public goods games through the K index [32], which represents the relation between the payoffs of all possible combinations of players' choices. Formally, $K = (R - P)/(T - S)$, where R denotes the payoff for joint cooperation (reward), P for joint defection (punishment), T for unilateral defection (temptation), and S for unilateral cooperation (sucker) (for additional information, see the electronic supplementary material). Accordingly, situations with a low K index pose greater conflict as they involve more temptation to defect and/or little differences in the outcomes resulting from joint cooperation and joint defection.

Along with the micro-structure of the interaction, the broader societal context has been considered another potential source of variation of gender differences in cooperative behaviour. Accordingly, gender differences can emerge across societies in response of the existing socio-ecological conditions. Recent empirical evidence is in line with this prediction, detecting substantial variation in the magnitude of gender differences in cooperation in a Prisoner's Dilemma setting across 10 societies [5] and 12 societies [4]. However, evidence about the societal factors underlying this variation is still limited. According to social role theory, gender roles are not established in a social vacuum, but they are deeply rooted in the division of labour in a given society [13]. The division of labour, in turn, is affected not only by sex-specific physical and reproductive attributes, but also by the demands posed by the social and economic environment [11]. Thus, it is possible to expect more cooperation from women in societies characterized by low levels of gender equality and economic development, as in these societies gender stereotypes might be more likely to be pervasive. To date, evidence supporting this prediction is

mixed. For example, in matrilinear societies in which traditional gender roles are reversed, women display more competitiveness than men in competitive tasks, compared to patriarchal societies that more closely resemble Western cultures gender roles [33]. On the other hand, greater gender differentiation in risk and social preferences (e.g. altruism, trust) has been observed in more egalitarian and developed countries [34] (referred to as the gender equality paradox). Also, other studies failed to detect any association between gender inequality and gender differences in cooperation across societies (e.g. [5]). Thus, it becomes relevant to test these predictions across highly standardized settings and a broad range of societies to contribute to the scientific understanding of whether and how gender differences in cooperation vary across societies.

In sum, the current meta-analysis tested the following set of hypotheses that have not yet been investigated using a meta-analytic approach:

- HP_1 : women cooperate more compared to men in social dilemmas;
- HP_2 : women, compared to men, cooperate more in situations with more conflicting interests;
- HP_3 : women, compared to men, cooperate more in societies characterized by lower gender equality;
- HP_4 : women, compared to men, cooperate more in societies characterized by lower economic development;

Finally, we pre-registered to conduct a conceptual replication testing several hypotheses from social role theory and evolutionary theories that were described and tested in a previous meta-analysis Balliet *et al.* [12], but with the addition of more recent studies [35], additional control for structural variables that have been recently found to be most strongly related to cooperation in social dilemma games [19], and through an alternative analytic approach. Specifically, besides analysing within-study gender differences in cooperation, we used a meta-regression model to test whether the gender composition of the sample could account for the observed cooperation across studies. These tests included the main effect of gender on cooperation, as well as the moderating effects of gender composition of the group, game iterations, group size, and year of data collection.

2. Methods

(a) Search for studies and inclusion criteria

Studies included in the analyses were retrieved through the Cooperation Databank (CoDa) [35]. Details of the systematic search that led to the annotation of the studies are reported in full in Spadaro *et al.* [35]. For the current work, we contacted corresponding authors of 44 articles retrieved from and annotated in CoDa and asked for the missing information to compute the effect size. Overall, 32% of the authors were able to provide the requested data, resulting in the inclusion of 18 effects. Pre-registration of hypotheses and analyses plan is reported on the OSF (<https://osf.io/m4t5r/>), together with the data and R scripts to perform the analyses. The data are also made openly available via CoDa: cooperationdatabank.org.

For studies to be included in the meta-analysis the following criteria had to be fulfilled:

1. Cooperation is assessed in a Prisoner's Dilemma, public goods or resource dilemma. Slight variations of these games were also included (e.g. asymmetrical payoff structure);

2. Cooperation is assessed exclusively in interactions with strangers;
3. Cooperation is provided over all trials of the game or during the first period;
4. Country of data collection is identifiable and unique for each study;
- 5a. Cooperation is reported separately for men and women (only relevant for analyses involving the gender-cooperation effect size);
- 5b. Gender composition of the sample is reported (only relevant for analyses involving logit-transformed cooperation rates).

These criteria resulted in the inclusion of 126 gender-cooperation effect sizes extracted from 121 studies conducted across 20 industrialized societies, and 972 study-level logit-transformed cooperation rates. Flowcharts detailing the outcomes of the search and inclusion criteria are reported in the electronic supplementary material.

(b) Coding of effect sizes

We conducted two sets of analyses using two measures of effect sizes to model the relationship between gender and cooperation. To capture standardized mean differences in cooperation between men and women, we used Cohen's d [36]. Based on our annotation, positive values of d indicate more cooperation from men, compared to women. We computed Cohen's d using means and standard deviations of contributions or withdrawals for games involving continuous choices, and proportion of cooperation and sample size for dichotomous choice games. To document the association between the gender composition of the sample and average cooperation in the game, we obtained a standardized logit-transformed measure of cooperation [37], computed differently for games with continuous and dichotomous choice types. See the electronic supplementary material for a complete description of the effect size and variance computation procedures. In the case of missing values in the standard deviation of contributions or withdrawals within a study, we imputed the median value of the coefficient of variation (CV) of all the studies reporting standard deviation (Prisoner's Dilemma and public goods games: $CV_{Mdn} = 0.490$, $k = 290$; resource dilemmas: $CV_{Mdn} = 0.353$, $k = 41$).

(c) Coding of variables

We included several relevant variables for differences between the economic games, including the following:

Gender composition of the group. In some of the studies, individuals interacted in groups with mixed gender composition (both men and women, $k = 58$), while in other studies groups were homogeneous with regard to gender ($k = 50$), or the study included both homogeneous and mixed groups ($k = 9$).

K index. We operationalized the degree of conflicting interests as K index [32], for both Prisoner's Dilemma and public goods games [30]. This index ranges from 0 to 1, with lower values indicating higher relative gains from defection over cooperation, resulting in a stronger conflict between self and collective interests in the game ($M = 0.46$, $Mdn = 0.40$, $s.d. = 0.21$).

Iterations. Participants could make one-shot decisions ($k = 41$), make repeated choice with the same partner(s) across multiple iterations of the game ($k = 82$), or both within the same game ($k = 3$).

Group size. Group size was indicated by the overall number of people affected by the choices in the game. Given the skewed distribution of the values (range = 2–50, $Mdn = 2$), group size was log-transformed prior to perform the analyses.

Dilemma type. Cooperation was assessed in three commonly used social dilemma paradigms. In Prisoner's Dilemmas ($k = 92$), participants are asked to decide independently whether to cooperate by transferring any portion x of their endowment e to the partner, knowing that x will be multiplied by a constant

m ($m > 1$) and added to the partner's endowment. In public goods games ($k = 28$), participants interact in groups of size N and each member decides how much to contribute to a group account, knowing that all contributions will be multiplied by m ($1 < m < N$) irrespective of individual contributions. In resource dilemmas ($k = 6$), participants interact in groups and each member decides how much to withdraw from a common pool, knowing that the amount will no longer be available to other group members and that the resource can be depleted, but can recover with reproduction rate $r > 1$.

Sanction. Some of the studies employed a sanction mechanism (either punishment or reward, $k = 3$), while most of the studies did not ($k = 122$), or manipulated this variable within the study ($k = 1$).

Discussion. Participants were allowed to communicate in some of the studies ($k = 12$), while in some other studies communication was explicitly forbidden ($k = 108$) or manipulated within the study ($k = 6$).

Additionally, we included sample characteristics such as *gender composition of the sample* (i.e. overall proportion of males in the sample of the study, $M = 0.50$, s.d. = 0.24), *year* (range = 1961–2017) and *country of data collection*. In case a study presented multiple values for a single study characteristic variable (e.g. if one or more study characteristics variables were manipulated within a study), we converted it in a single value for numeric and categorical variables by using either the median or annotating a 'mixed' level. Inter-rater agreement for the annotation of the study characteristics was estimated to be medium-to-high ($0.68 < \text{Krippendorff's } \alpha < 0.97$) (see Spadaro *et al.* [35] for details).

We included additional variables at the societal level to test hypotheses about whether the relationship between gender and cooperation is related to societal differences in gender equality and economic development:

Gender equality. Following Falk & Hermlé [34] we selected four indices of gender equality: the Global Gender Gap Index [1], the Gender Inequality Index [38] (reversed), ratio of female to male labour force participation rates [39] (reversed), and the number of years since women's suffrage [40].

Economic development. We used World Bank's measures of GDP *per capita*, GNI, and Gini coefficient (reversed) to operationalize societies' economic development [41].

Indicators were associated with the effect sizes based on the specific country and year of data collection of each study, so that if multiple time points are available for the same indicator within a society, we matched the effect with the closest time point to the year of data collection. To prevent multicollinearity and increase statistical power, analyses including gender equality and economic development indicators were performed using principal components (and not single indicators) extracted from the two sets of indicators (see electronic supplementary material for more detail).

(d) Analytic strategy

To test HP_1 , we took a twofolded approach. We analysed estimates of the within-study differences in cooperation between men and women (i.e. Cohen's d) by using a random effects meta-analytic model with REML estimation [37]. Additionally, we tested whether the gender composition of the sample could account for observed variation in mean levels of cooperation across studies using a meta-regression model. To test HP_2 , we complemented both models with the K index variable and an interaction term crossing the K index with the gender composition of the sample, respectively.

We then added two sets of moderators to all models to replicate previous meta-analytic evidence [12], and to control for additional contextual variables that differed between studies (i.e. the provision of sanctions and the opportunity to communicate). Together with the K index, these variables have been

shown to be the strongest predictors of cooperative behaviour in economic games settings, although no specific hypotheses were forwarded in relation to gender differences *per se* [19].

To test $HP_{3,4}$, we extended the model analysing within-study differences in cooperation by including the society in which the study has been conducted as random effects [42]. We first tested the significance of the between-society variance component through a one-sided log-likelihood ratio test that compared the fit of an intercept-only model to the fit of a model in which there were no predictors as well, but where between-society variance was manually constrained to zero. Then, we included the societal-level indicators of gender equality and economic development as moderators of the effect size.

All analyses were conducted in R [43] using the *metafor* package [44]. Following the pre-registration, we used multiple imputation of missing data for models whose variables presented missing information using the *mice* package [45] (number of missing cases per variable is reported in electronic supplementary material, table S5). The results of analyses obtained through listwise deletion are reported in the electronic supplementary material as a robustness check, and fully replicated the pattern of findings observed using the imputed dataset.

3. Results

(a) Mean gender differences in cooperative behaviour

First, to test HP_1 , we meta-analysed the mean differences in cooperative behaviour between men and women in social dilemmas studies to test whether there was any overall gender difference in cooperation. To do so, we fitted an intercept-only meta-analytic model ($k = 126$). Results from the random effects model showed no significant differences in cooperation between men and women ($d = 0.011$, 95% CI $[-0.039, 0.060]$, 90% prediction interval $[-0.354, 0.375]$). Moreover, there is heterogeneity in the effect size distribution ($T = 0.185$, $T^2 = 0.034$). Most of this heterogeneity can be attributed to between-study differences ($I^2 = 0.509$). Egger's regression test for publication bias revealed a non-significant association between the magnitude of the gender effect size and the standard error of the estimate ($z = 1.279$, $p = 0.201$), suggesting little evidence for bias in the analysed data. The estimated overall effect did not change when applying less restrictive inclusion criteria that included additional effect sizes from studies reporting cooperation in other periods of the game ($k = 7$), ($d = 0.011$, 95% CI $[-0.037, 0.060]$, 90% prediction interval $[-0.348, 0.370]$).

To replicate and extend previous research on the contextual factors that can moderate the relationship between gender and cooperation, we added two sets of moderators to the meta-analytic model. First, as done in Balliet *et al.* [12] we included gender composition of the group, iterations, group size, dilemma type and year of data collection as study-level moderators. The annotation of study-level moderators and effect size computed for each study included in the meta-analysis is reported in electronic supplementary material, table S5. In contrast to previous meta-analytic evidence, none of the study-level variables was significantly associated with the gender effect size (p -values ≥ 0.363) (electronic supplementary material, table S6). Then, as a robustness check, we included an additional set of moderators (i.e. discussion, sanction) to control for the structural variables that have been found to be most strongly related to cooperation in social dilemma games [19]. Again, we did not find any significant association between the study-level

moderators and the gender effect size (p -values ≥ 0.153) (electronic supplementary material, table S6).

We then tested the hypothesis HP_2 , according to which differences in cooperation between men and women might be moderated by the degree of conflicting interest in the situation. To do so, we selected effects from studies involving a Prisoner's Dilemma or a public good game for which a K index is computable ($k = 108$) and included the K index as a predictor in the model, with discussion and sanction as further control variables. In contrast to the predictions, differences in cooperation between men and women did not change according to the degree of conflicting interests in the situation ($p = 0.512$). This pattern of null findings was mostly replicated when running the analysis without performing imputation of missing data (results are provided in the electronic supplementary material).

To examine whether the broader cultural and institutional context affect gender differences in cooperation, we extended the structure of the model to a multi-level structure, and nested studies within countries. Results of the log-likelihood ratio test showed that the model in which between-society variance was freely estimated did not have a significantly better fit than the model in which between-society variance was not modelled (LRT < 0.001 , $p = 1.000$). This finding suggests that there is no significant between-society variation in gender differences in cooperation.

Then, to test our hypotheses $HP_{3,4}$ about how gender differences in cooperation relate to societal gender equality and economic development, we ran two multi-level meta-analytic models including these societal-level indicators as additional predictors. For each model, the indicators were first aggregated in principal components to avoid collinearity, and the presence of sanctions and communication opportunities were included as further control variables. Our analyses revealed that neither societal gender equality (p -values ≥ 0.078) and economic development (p -values ≥ 0.060) were significantly associated with the gender differences in cooperation (electronic supplementary material, table S7).

(b) Variation in mean levels of cooperation across studies and gender composition of the sample

We first meta-analysed mean levels of cooperation observed in social dilemmas studies by fitting an intercept-only meta-regression model that predicted logit-transformed cooperation. Overall, across all studies included in the meta-analysis, the mean cooperation rate was 0.505 (95% CI [0.494, 0.516], 90% PI [0.224, 0.783]) ($k = 972$), meaning that participants contributed on average half of their endowment in the game. Moreover, there is variation in the distribution of cooperation estimates ($T = 0.644$, $T^2 = 0.415$), that is greater than it would be expected by chance alone ($Q(971) = 56\,656.296$, $p < 0.001$). Such variation in the distribution of cooperation estimates can be mostly attributed to between-study differences ($I^2 = 0.983$). Egger's regression test for publication bias suggests an absence of bias in the analysed data ($z = 0.616$, $p = 0.538$).

We then tested HP_1 using an alternative methodological approach and examined whether the gender composition of the sample could account for the observed variation in cooperation in the studies. To do so, we ran a meta-regression model that included the gender composition of the sample and two study characteristics variables that have been shown to be strongly associated with cooperation (i.e.

discussion and sanctions) as predictors. Gender composition of the sample (i.e. overall proportion of males) was negatively and significantly associated with cooperation ($b = -0.401$, $p < 0.001$), suggesting that higher cooperation was observed in studies with a prevalence of women ($\Delta R^2 = 0.0199$).

To test pre-registered HP_2 , positing a potential interaction between the gender composition of the sample and the degree of conflicting interests in predicting cooperation, we added an interaction term between gender composition and the K index. As the K index can be only calculated for Prisoner's Dilemma games and public good games, we only selected these two types of dilemmas in the remaining analyses ($k = 907$). In contrast to the pre-registered hypothesis, findings revealed a non-significant interaction between the K index and gender composition of the sample ($p = 0.920$). Moreover, gender composition of the sample became non-significant while controlling for the interaction ($p = 0.233$) (see electronic supplementary material, table S8).

4. Discussion

This meta-analysis examined empirical studies on cooperation using social dilemma paradigms to answer questions about the relationship between participants' gender and cooperative behaviour. Specifically, we tested whether women are overall more cooperative than men, and novel pre-registered hypotheses about the moderating role of contextual factors such as the degree of conflict in the situation, and societal adherence to canonical gender roles and economic development. Overall, we found little-to-no evidence for gender differences in cooperation and no support for the additional moderation hypotheses. Below, we discuss these findings, their limitations, and suggest some potential directions for future research.

In line with previous meta-analytic evidence [12], we found no within-study differences in cooperation between men and women. Men and women displayed comparable levels of cooperation in Prisoner's Dilemmas, public goods dilemmas and resource dilemma games ($k = 126$, $d = 0.011$). However, we did find a significant association between overall gender composition of the sample and the mean levels of cooperation across 972 studies. This result suggests that there is higher cooperation in studies with a higher prevalence of women. Although this latter analysis benefits from a large number of studies, societies, and experimental settings, we should interpret these findings with caution based on (a) potential methodological confounds related to changes in samples over time, and (b) conflict with existing evidence. In fact, over the last 60 years, cooperation in studies using economic games has increased over time [46], and so did the inclusion of women in the experimental samples [47]. In our data, year of data collection is both positively correlated with logit-transformed cooperation rates ($r = 0.17$, $p < 0.001$) and negatively associated with proportion of men in the sample ($r = -0.27$, $p < 0.001$). Although the association of gender composition of the sample and cooperation remains significant while controlling for year of data collection ($b = -0.286$, $p = 0.002$), we could not rule out that temporal trends in methodologies could account for the observed significant association. In addition, this analysis does not replicate the result of a similar analysis using a broader set of studies ($N = 1527$) and that controls for a greater number

of study characteristics (e.g. mean age of the sample, discipline of study, symmetry, deception) [19]. Considering these concerns, we conclude that we do not find compelling evidence in support of gender differences in cooperation.

We further tested whether women cooperate more than men in situations involving greater conflict of interests. The findings did not provide support for this prediction, either examining whether the degree of conflict (i) moderated within-studies gender differences in cooperation or (ii) interacted with gender composition of the sample to predict mean levels of cooperation across studies. Given that situations with higher conflict of interests involve more risk of exploitation, these null findings can also inform research investigating whether gender differences in cooperation relate to gender-specific attitudes toward risks [5,34,48]. The severity of conflict in the meta-analysis was operationalized using the payoff structure (i.e. the *K* index, [32]) of games that afford the potential for exploitation [30]. Although this approach had the advantage of evaluating the moderation of conflict within situations that had a similar incentive structure, the studies included in the meta-analysis presented little variation in the *K* index. In fact, the *K* index ranged from 0.20 to 0.40 for 46% of the studies ($M=0.46$, $Mdn=0.40$, $s.d.=0.21$). Although this is in line with what is observed across all studies in CoDa (e.g. 39% ranging between 0.20 and 0.40, [49]), variation within the *K* index might be too small to detect any differential responses to stress or emergence of canonical gender roles to result in gender differences in cooperative responses. An alternative way to test this hypothesis could be to examine gender differences in cooperative behaviour across game situations with weak or strong exploitation components (e.g. as done in [31] by comparing behaviour in a ultimatum game and a Prisoner's Dilemma). In addition, the type of conflict of interests faced in situations resembling a Prisoner's Dilemma structure might only be a small fraction of the situations experienced in daily life [50], and Prisoner's Dilemma situations might provide situational cues of the potential of exploitation that could affect the occurrence of gender-typical behaviour [27]. A promising avenue for future research might be then to examine cooperative behaviour by systematically varying other relevant situational features, such as introducing the possibility to benefit others through one's competitive behaviour (e.g. socially oriented incentives, [51]) and the information about the interaction partner (e.g. anonymity, [5]). The identification of additional contextual features can contribute to a more comprehensive understanding of the mixed patterns of findings on gender differences in cooperation.

In addition, in the present meta-analysis, gender differences in cooperation did not vary across societies. This evidence is in line with studies showing little evidence for cross-cultural variation of gender differences in prosocial behaviour in children from both industrialized and small-scale societies [52]. Here, studies conducted in societies at different levels of gender equality and economic development displayed very similar cooperation by men and women. Although differences would be expected in light of social role theory [13,27], these findings are in line with recent empirical evidence showing that gender inequality was not associated with differences in magnitude of gender differences in cooperative behaviour in a Prisoner's Dilemma in 10 countries [5]. Compared to Dorrough & Glöckner's study [5], our meta-analysis included a broader range of

countries and societal indicators, such as ratio of female to male labour force participation rates [39] and the number of years since women's suffrage [40]. However, none of these seven indicators was significantly correlated to mean differences in cooperation across studies (p -values ≥ 0.292). These findings are consistent with recent meta-analytic work showing no evidence for cross-cultural variation in cooperation more broadly [53]. It is worth mentioning, however, that despite our effort to obtain more studies (e.g. through direct requests to authors), findings from more recent cross-cultural studies detecting gender differences in cooperation could not be included, since these studies were not yet annotated in CoDa (e.g. [4,5]). As such, our meta-analysis has low statistical power to detect variation across societies, due to the limited number of effects available for each society. Moreover, the included studies mostly comprised WEIRD samples [54] and might not be representative of the actual cross-cultural variation in cooperative behaviour. For a more comprehensive analysis, we encourage future work to more systematically disclose information about cooperation displayed by men and women, or to provide this information retrospectively for previously published studies (e.g. through CoDa [35]). At present, however, these limitations might impact the reliability of variance observed at the highest level of the model. More research is needed to replicate our findings with a broader set of societies and observations.

Last, we tested whether other features of the interaction context moderated gender differences in cooperation to provide a conceptual replication of findings from previous meta-analytic work [12]. We did replicate that gender is not associated with cooperation and that group size and year of data collection do not significantly moderate the gender effects after controlling for several study characteristics. However, we found no support for the moderation hypotheses related to gender composition of the group, group size, iterations, and year of data collection, as none of these variables were significantly associated with the magnitude of gender differences in cooperation. These different patterns of findings might be due to the way primary studies have been selected in the present work, namely the inclusion of more recent studies (2010–2017, $k=37$), and the adoption of stricter inclusion criteria (e.g. matrix games not classifiable as Prisoner's Dilemma and public goods games were not included, and so did studies involving interactions among acquaintances). It is worth noting that our goal was not to perform an exact replication of previous work. Nevertheless, the conceptual replication of the main effect provides even stronger evidence that there is no main overall association between gender and cooperation. Furthermore, the lack of moderation of the association between gender and cooperation suggests that these moderation effects are not very robust to variations in the data selection and analytic techniques—and so should be interpreted with caution. The fact that the moderating effects were not robust to these adjustments suggest that even small variations of the context can be crucial to elicit (or not) gender differences in cooperation. For example, the present meta-analysis included additional studies from more recent years, and there have been changes in samples and methods in the literature over time, such as more recent studies having (a) a greater percentage of women, (b) fewer student samples, and (c) a stronger conflict of interests (i.e. lower *K* value) [47]. Future studies might consider to experimentally manipulate the situational

features of interest, such as gender composition of the group (e.g. [55]) and conflict of interests (e.g. [56]), to provide a further test of these moderating hypotheses.

5. Conclusion

Decades of theory have attempted to explain the potential existence of gender differences in prosocial and cooperative behaviour. Since the earliest studies on social dilemmas (e.g. [57]), literature presented mixed findings in support of gender differences in cooperative behaviour [55], suggesting that they can emerge in response to specific features of the context and of the interaction (e.g. [9]). Inflated assumptions about gender differences, however, can pose societal costs [58] and continue to have ramifications in terms of reduced opportunities for women to access, contribute to, and manage public goods and common resources in modern societies. Thus, it becomes important to enhance our scientific understanding of the phenomenon. We meta-analysed 56 years of studies that used a highly standardized experimental setting (i.e. experimental social dilemmas) and found very limited evidence for any observed differences in cooperation

between men and women. This finding replicated previous meta-analytic evidence. Altogether, these results contribute to our understanding of the generalizability of gender differences across different contexts and supports the idea that women's cooperation is rather flexible and adapts to different contexts [59].

Data accessibility. Pre-registration of hypotheses and analyses plan is reported on the OSF (<https://osf.io/m4t5r/>), together with the data and R scripts to perform the analyses. The data are also made openly available via CoDa: cooperationdatabank.org.

The data are provided in the electronic supplementary material [60].

Authors' contributions. G.S.: conceptualization, formal analysis, investigation, methodology, writing—original draft, writing—review and editing; S.J.: conceptualization, investigation, methodology, visualization, writing—review and editing; D.B.: conceptualization, funding acquisition, methodology, supervision, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. This research was supported by an ERC-Starting (grant no. 635356) and an ERC-Consolidator (grant no. 864519) awarded to D.B.

References

- Economic Forum. 2020 Global Gender Gap Report 2020. See https://www3.weforum.org/docs/WEF_GGGR_2020.pdf.
- Eagly AH. 2009 The his and hers of prosocial behavior: an examination of the social psychology of gender. *Am. Psychol.* **64**, 644–658. (doi:10.1037/0003-066X.64.8.644)
- Sell J, Kuipers KJ. 2009 A structural social psychological view of gender differences in cooperation. *Sex Roles* **61**, 317–324. (doi:10.1007/S11199-009-9597-5)
- Dorough AR, Glöckner A. 2019 A cross-national analysis of sex differences in Prisoner's Dilemma games. *Br. J. Soc. Psychol.* **58**, 225–240. (doi:10.1111/BJSO.12287)
- Dorough AR, Glöckner A. 2021 Sex differences concerning prosocial behavior in social dilemmas are (partially) mediated by risk preferences but not social preferences: an in-depth analysis across 10 countries. *Soc. Psychol. (Gott)* **52**, 61–74. (doi:10.1027/1864-9335/A000434)
- Romano A, Sutter M, Liu JH, Yamagishi T, Balliet D. 2021 National parochialism is ubiquitous across 42 nations around the world. *Nat. Commun.* **12**, 4456. (doi:10.1038/s41467-021-24787-1)
- Walters AE, Stuhlmacher AF, Meyer LL. 1998 Gender and negotiator competitiveness: a meta-analysis. *Organ. Behav. Hum. Decis. Process.* **76**, 1–29. (doi:10.1006/OBHD.1998.2797)
- Ben-Ner A, Kong F, Putterman L. 2004 Share and share alike? Gender-pairing, personality, and cognitive ability as determinants of giving. *J. Econ. Psychol.* **25**, 581–589. (doi:10.1016/S0167-4870(03)00065-5)
- Simpson B, Van Vugt M. 2009 Sex differences in cooperation: integrating the evolutionary and social psychological perspectives. *Adv. Gr. Process.* **26**, 81–103. (doi:10.1108/S0882-6145(2009)0000026007)
- Eagly AH, Wood W. 2012 Social role theory. In *Handbook of theories of social psychology: volume 2* (eds PAM Van Lange, AW Kruglanski, ET Higgins), pp. 458–476. Thousand Oaks, CA: SAGE Publications Ltd.
- Wood W, Eagly AH. 2002 A cross-cultural analysis of the behavior of women and men: implications for the origins of sex differences. *Psychol. Bull.* **128**, 699–727. (doi:10.1037//0033-2909.128.5.699)
- Balliet D, Li NP, Macfarlan SJ, Van Vugt M. 2011 Sex differences in cooperation: a meta-analytic review of social dilemmas. *Psychol. Bull.* **137**, 881–909. (doi:10.1037/a0025354)
- Wood W, Eagly A. 2010 Gender. In *Handbook of social psychology* (eds ST Fiske, DT Gilbert, G Lindzey), pp. 629–667. New York, NY: John Wiley & Sons, Inc.
- Buss DM. 1995 Psychological sex differences: origins through sexual selection. *Am. Psychol.* **50**, 164–168. (doi:10.1037/0003-066X.50.3.164)
- Van Vugt M. 2009 Sex differences in intergroup competition, aggression, and warfare: the male warrior hypothesis. *Ann. N. Y. Acad. Sci.* **1167**, 124–134. (doi:10.1111/J.1749-6632.2009.04539.X)
- Kenrick DT, Li NP, Butner J. 2003 Dynamical evolutionary psychology: individual decision rules and emergent social norms. *Psychol. Rev.* **110**, 3–28. (doi:10.1037/0033-295X.110.1.3)
- Geary DC, Byrd-Craven J, Hoard MK, Vigil J, Numtee C. 2003 Evolution and development of boys' social behavior. *Dev. Rev.* **23**, 444–470. (doi:10.1016/J.DR.2003.08.001)
- Benenson JF, Markovits H, Fitzgerald C, Geoffrey D, Flemming J, Kahlenberg SM, Wrangham RW. 2009 Males' greater tolerance of same-sex peers. *Psychol. Sci.* **20**, 184–190. (doi:10.1111/j.1467-9280.2009.02269.x)
- Jin S, Spadaro G, Balliet D. In preparation. Institutions and cooperation: a meta-analysis of rules in social dilemmas.
- Balliet D, Van Lange PAM. 2013 Trust, conflict, and cooperation: a meta-analysis. *Psychol. Bull.* **139**, 1090–1112. (doi:10.1037/a0030939)
- Taylor SE, Klein LC, Lewis BP, Gruenewald TL, Gurung RAR, Updegraff JA. 2000 Biobehavioral responses to stress in females: tend-and-befriend, not fight-or-flight. *Psychol. Rev.* **107**, 411–429. (doi:10.1037/0033-295X.107.3.411)
- Cannon WB. 1932 *The wisdom of the body*. New York, NY: Norton.
- Taylor SE. 2002 *The tending instinct: how nurturing is essential to who we are and how we live*. New York, NY: Holt.
- Orbell J, Dawes R, Schwartz-Shea P. 1994 Trust, social categories, and individuals: the case of gender. *Motiv. Emot.* **18**, 109–128. (doi:10.1007/BF02249396)
- Greig F, Bohnet I. 2009 Exploring gendered behavior in the field with experiments: why public goods are provided by women in a Nairobi slum. *J. Econ. Behav. Organ.* **70**, 1–9. (doi:10.1016/J.JEBO.2008.12.006)
- Brañas-Garza P, Capraro V, Rascón-Ramírez E. 2018 Gender differences in altruism on Mechanical Turk: expectations and actual behaviour. *Econ. Lett.* **170**, 19–23. (doi:10.1016/J.ECONLET.2018.05.022)
- Wood W, Eagly AH. 2009 Gender identity. In *Handbook of individual differences* (eds M Leary, R Hoyle), pp. 109–128. New York, NY: Guilford Press.
- Balliet D, Wu J, De Dreu CKW. 2014 Ingroup favoritism in cooperation: a meta-analysis. *Psychol. Bull.* **140**, 1556–1581. (doi:10.1037/a0037737)

29. Schopler J, Insko CA, Wieselquist J, Pemberton M, Witcher B, Kozar R, Roddenberry C, Wildschut T. 2001 When groups are more competitive than individuals: The domain of the discontinuity effect. *J. Pers. Soc. Psychol.* **80**, 632–644. (doi:10.1037//0022-3514.80.4.632)
30. Thielmann I, Spadaro G, Balliet D. 2020 Personality and prosocial behavior: a theoretical framework and meta-analysis. *Psychol. Bull.* **146**, 30–90. (doi:10.1037/bul0000217)
31. Nickels N, Kubicki K, Maestriperi D. 2017 Sex differences in the effects of psychosocial stress on cooperative and prosocial behavior: evidence for ‘flight or fight’ in males and ‘tend and befriend’ in females. *Adapt. Hum. Behav. Physiol.* **3**, 171–183. (doi:10.1007/S40750-017-0062-3)
32. Rapoport A, Chammah AM. 1965 *Prisoner’s Dilemma: a study in conflict and cooperation*. Ann Arbor, MI: University of Michigan Press.
33. Gneezy U, Leonard KL, List JA. 2009 Gender differences in competition: evidence from a matrilineal and a patriarchal society. *Econometrica* **77**, 1637–1664. (doi:10.3982/ECTA6690)
34. Falk A, Hermlé J. 2018 Relationship of gender differences in preferences to economic development and gender equality. *Science* **362**, eaas9899. (doi:10.1126/SCIENCE.AAS9899)
35. Spadaro G, Tiddi I, Columbus S, Jin S, ten Teije AACM, CoDa Team, Balliet D. 2022 The Cooperation Databank: machine-readable science accelerates research synthesis. *Perspect. Psychol. Sci.* **17**, 1472–1489. (doi:10.1177/17456916211053319)
36. Cohen J. 1988 *Statistical power analysis for the behavioral sciences*. New York, NY: Routledge Academic.
37. Lipsey MW, Wilson DB. 2000 *Practical meta-analysis*. Thousand Oaks, CA: Sage.
38. United Nations. 2020 Human Development Reports: Gender Inequality Index (GII). See <https://hdr.undp.org/data-center/thematic-composite-indices/gender-inequality-index#/indicies/GII>.
39. World Bank. 2019 Ratio of female to male force participation rate. See <https://data.worldbank.org/indicator/SL.TLF.CACT.FM.ZS>.
40. Inter-Parliamentary Union. 2005 Women’s suffrage: a world chronology of the recognition of women’s rights to vote and to stand for election. See <http://archive.ipu.org/wmn-e/suffrage.htm>.
41. World Bank. 2018 World Development Indicators. See <https://databank.worldbank.org/source/world-development-indicators>.
42. Van den Noortgate W, López-López JA, Marín-Martínez F, Sánchez-Meca J. 2013 Three-level meta-analysis of dependent effect sizes. *Behav. Res. Methods* **45**, 576–594. (doi:10.3758/s13428-012-0261-6)
43. R Core Team. 2019. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. See <http://www.R-project.org/>.
44. Viechtbauer W. 2010 Conducting meta-analyses in R with the metafor Package. *J. Stat. Softw.* **36**, 1–48. (doi:10.18637/jss.v036.i03)
45. van Buuren S. 2018 *Flexible imputation of missing data*, 2nd edn. New York, NY: Chapman and Hall.
46. Yuan M, Spadaro G, Jin S, Wu J, Kou Y, Van Lange PAM, Balliet D. 2022 Did cooperation among strangers decline in the United States? A cross-temporal meta-analysis of social dilemmas (1956–2017). *Psychol. Bull.* **148**, 129–157. (doi:10.1037/bul0000363)
47. Balliet D, Spadaro G, Markovitch B, Beek W. 2020 How did cooperation research change over time? See <https://cooperationdatabank.org/data-stories/how-did-cooperation-research-change-over-time/>.
48. Charness G, Gneezy U. 2012 Strong evidence for gender differences in risk taking. *J. Econ. Behav. Organ.* **83**, 50–58. (doi:10.1016/J.JEBO.2011.06.007)
49. Balliet D, Spadaro G, Markovitch B, Beek W. 2020 What’s in the Databank? See <https://cooperationdatabank.org/data-stories/whats-in-the-databank/>.
50. Columbus S, Molho C, Righetti F, Balliet D. 2020 Interdependence and cooperation in daily life. *J. Pers. Soc. Psychol.* **120**, 626–650. (doi:10.1037/pspi0000253)
51. Cassar A, Ridgion M. 2022 Sustaining the potential for cooperation as female competitive strategy. *Phil. Trans. R. Soc. B* **378**, 20210440. (doi:10.1098/rstb.2021.0440)
52. House B, Silk JB, McAuliffe K. 2022 No strong evidence for universal gender differences in the development of cooperative behaviour across societies. *Phil. Trans. R. Soc. B* **378**, 20210439. (doi:10.1098/rstb.2021.0439)
53. Spadaro G *et al.* 2022 Cross-cultural variation in cooperation: a meta-analysis. *J. Pers. Soc. Psychol.* **123**, 1024–1088. (doi:10.1037/pspi0000389)
54. Henrich J, Heine SJ, Norenzayan A. 2010 The weirdest people in the world? *Behav. Brain Sci.* **33**, 61–83. (doi:10.1017/S0140525X0999152X)
55. Sell J, Griffith WI, Wilson RK. 1993 Are women more cooperative than men in social dilemmas? *Soc. Psychol. Q.* **56**, 211. (doi:10.2307/2786779)
56. Vlaev I, Chater N. 2006 Game relativity: how context influences strategic decision making. *J. Exp. Psychol. Learn. Mem. Cogn.* **32**, 131–149. (doi:10.1037/0278-7393.32.1.131)
57. Rapoport A, Chammah AM. 1965 Sex differences in factors contributing to the level of cooperation in the Prisoner’s Dilemma game. *J. Pers. Soc. Psychol.* **2**, 831–838. (doi:10.1037/H0022678)
58. Hyde JS. 2005 The gender similarities hypothesis. *Am. Psychol.* **60**, 581–592. (doi:10.1037/0003-066X.60.6.581)
59. Fox SA, Scelza B, Silk J, Kramer KL. 2022 New perspectives on the evolution of women’s cooperation. *Phil. Trans. R. Soc. B* **378**, 20210424. (doi:10.1098/rstb.2021.0424)
60. Spadaro G, Jin S, Balliet D. 2022 Gender differences in cooperation across 20 societies: a meta-analysis. Figshare. (doi:10.6084/m9.figshare.c.6249964)