Psychol Stud (October–December 2020) 65(4):360–369
https://doi.org/10.1007/s12646-020-00573-9

REVIEW ARTICLE

Gender Differences in Solving Moral Dilemmas: Emotional Engagement, Care and Utilitarian Orientation

P. Cordellieri[1] · M. Boccia[1,2] · L. Piccardi[1,2] · D. Kormakova[1] · L. V. Stoica[1] · F. Ferlazzo[1] · C. Guariglia[1,2] · A. M. Giannini[1]

Received: 27 June 2019 / Accepted: 2 October 2020 / Published online: 25 November 2020

Abstract Moral sense is important for determining human behaviour. Moral sense becomes crucial in operational environments in which choices must be made that have complex moral implications in highly stressful situations. Behavioural and neuroimaging findings have shown the existence of gender-related differences in moral reasoning. The present study aimed to investigate whether gender affects moral reasoning and emotional state. We also investigated whether empathy, decision-making and emotional regulation strategies had a role in determining gender differences in solving moral dilemmas. We found that moral judgements and emotional engagement were significantly different. Women were less prone than men to accept a moral violation, such as killing someone to save their own lives and the lives of others. Furthermore, women were more emotionally involved and experienced dysphoric emotions more often than men. Our results shed light upon the mechanisms that affect moral reasoning and determine gender differences in solving moral dilemmas.

Keywords Gender moral difference · Moral judgment · Moral dilemmas · Moral justice

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s12646-020-00573-9) contains supplementary material, which is available to authorized users.

✉ P. Cordellieri
pierluigi.cordellieri@gmail.com

[1] Department of Psychology, 'Sapienza' University of Rome, Rome, Italy

[2] Neuropsychology Unit, IRCCS Fondazione Santa Lucia, Rome, Italy

Introduction

The existence of gender differences in moral judgement is a very controversial and highly debated issue, with different theoretical and methodological orientations. Kohlberg (1964a, b) proposed an explanation of moral judgement based primarily on the fundamental role of reason. He argued that the development of moral reasoning takes place in an invariant sequence of universal stages (*preconventional, conventional and postconventional*) in all cultures. According to Kohlberg's theory, progressive cognitive development in increasingly complex structures involves a series of qualitative transformations in understanding and resolving moral dilemmas, including a sense of equity and justice, which is considered to be the basic principle underlying moral behaviour. By using structured interviews aimed at solving moral dilemmas (Colby & Kohlberg, 1987), Kohlberg showed that men are more oriented towards considering universal rights and principles than women, almost assuming an ethical superiority in men.

This finding was disputed by Gilligan (1982), who considered the principles of equity and justice exclusively in relation to moral judgement, omitting the interpersonal and affective dimension (such as taking care of others). Men solve moral dilemmas in a rational way, that is, by respecting law and order. In contrast, women are driven by emotion, empathy and care for others. These differences led Gilligan to describe two divergent approaches to morality. The first consists of a care-based orientation, more prone in women, which emphasises interpersonal relationships and is guided by social emotions, including empathy and altruism. The second approach refers to a justice-based orientation, more prone in men, which emphasises maintaining order and adhering to rules.

Our goal, in this study, is to understand whether there are gender differences in moral judgment, particularly with regard to the emotional processes underlying this process. The question of gender differences is not yet clear, although there is a lot of research in this field.

Some studies confirmed the main use of reason (based on the categories of justice) in men and the principles of care in women (Gilligan & Attanucci, 1988; Yacker & Weinberg, 1990; Indick et al., 2000; Aldrich & Kage, 2003); others found gender-related differences only in the second case (Garmon et al., 1996; Gibbs et al., 1984; Wark & Krebs, 1966); others found no significant differences in the use of either mode (Beal et al., 1997; Friedman et al., 1981; Walker et al., 1987). The scientific debate that developed from these premises assumed two distinct orientations in the context of moral psychology: one orientation was aimed at supporting the existence of gender-related moral guidelines (Aldrich & Kage, 2003; Bjorklund, 2003; Gump et al., 2000; Indick et al., 2000), and the other orientation was aimed at establishing their substantial experimental unfoundedness (Brabeck & Shore, 2003; Jaffee & Hyde, 2000). However, this issue is still unresolved.

The neurosciences have made new and significant contributions to studying the processes of moral judgement. In particular, the neural basis of moral cognition and emotion have been investigated using neuroimaging methods (Greene & Haidt, 2002; Greene et al., 2004; Moll et al., 2005; Harenski & Hamann 2006). Greene and colleagues distinguished between *personal* and *impersonal* moral violations and judgements based on the dual-process theory (Greene et al., 2001; Greene & Haidt, 2002; Greene et al., 2004). Personal moral dilemmas require that the agent is directly involved in the production of the harm, as in cases where it is decided to inflict damage to gain a wider moral advantage. For example, deciding to kill someone to save more people, or to find a wounded animal and face the decision of ending the animal's life to end its suffering. Such decisions imply considerable emotional involvement. Impersonal moral dilemmas are those in which the agent is only indirectly involved in the process that results in the harm. What makes a dilemma impersonal is that the agent just initiates a process that, through its own dynamics, ends up causing the harm. For example, saving some people, but as a result of my action, someone else will die. Impersonal moral judgments are usually driven by in-depth cognitive activity. Different brain regions were found to be engaged in solving *personal* and *impersonal* moral dilemmas. The brain regions associated with emotion and social cognition (the medial prefrontal cortex, posterior cingulate cortex and precuneus, superior temporal sulcus and temporo-parietal junction) exhibited increased activity when participants considered personal moral dilemmas, and *cognitive* brain regions associated with abstract reasoning and problem-solving (the middle frontal gyrus and parietal lobe) exhibited increased activity when participants considered impersonal moral dilemmas (Greene et al., 2001; Boccia et al., 2017a). Although several criticisms have been levelled at the distinction between personal and impersonal moral dilemmas (Mikhail, 2007; McGuire et al., 2009; Boccia et al., 2017b), it must be acknowledged that Greene and collaborators were instrumental in clarifying the relationship between cognitive and emotional processes in moral judgement. In fact, Greene showed the existence of two different neural courses (Greene, 2009), one course driven by automatic emotional responses evoked in cases of *deontological* moral judgement (disapproving the killing of one person to save several others), and the other course mainly controlled by cognitive processes in the case of *utilitarian* judgement (approving the killing of one person to save the others).

Fumagalli et al. (2010), who embraced Greene's theorisation about personal and impersonal moral reasoning, recently found that responses to personal moral dilemmas differ specifically and selectively in men and women. In particular, men make more pragmatic choices and are less concerned than women about harming others. In contrast, women are more reluctant than men to make decisions that inflict physical or moral pain on someone (non-utilitarian response). Recently, gender differences were also found to be in implicit association with the concepts of justice and care in the Implicit Association Test (IAT; Agerström et al., 2011). The hypothesis of two different approaches to solving moral dilemmas was also confirmed in neuroimaging studies. For example, Harenski et al. (2008) found that women showed a positive modulation of violation severity ratings in the posterior cingulate cortex and the anterior insula. The role of these regions has been demonstrated in empathic processes and in problem-solving that calls for care-oriented situations (Botvinik et al., 2005; Singer et al., 2006; Robertson et al., 2007). Furthermore, compared with women, men showed a stronger modulatory relationship between inferior parietal activity and moral rating. These results suggest that women and men use different moral strategies when they are engaged in moral reasoning. In particular, women mainly rely on brain structures related to emotion and social cognition (the posterior cingulate cortex and insula), and men rely mainly on brain regions related to non-social cognition (the inferior parietal cortex) (Greene & Haidt, 2002).

Although it has been demonstrated that there are gender differences in several areas, such as in empathic ability (Baron-Cohen & Wheelwright, 2004), and other functions that have a causal role in moral judgement, such as emotions (Aleman & Swart, 2008; Harenski et al., 2008; Hareli et al., 2009; Verde et al., 2013), it is not yet clear

whether men and women face moral judgements with different cognitive and emotional patterns.

In this study, we aimed to determine whether men and women adopt different strategies in making moral choices. To pursue this aim, we asked our participants to evaluate 4 different dilemmas. The empathy, decision-making and emotional regulation strategies were also evaluated. After the participants responded to the questions concerning each dilemma, their self-assessment of their emotional state was also obtained. We hypothesised the presence of a difference between men and women in moral decision-making, particularly with regard to emotional processes. Specifically, we hypothesised that the choices of men are oriented more towards a pragmatic or utilitarian choice beyond putting others in a state of risk, danger or harm. In contrast, we hypothesised that the responses of women are more oriented towards ethics of care, supported by greater emotional salience and empathic sensitivity, with a perspective that is more often deontological ('not murder', 'not injure', 'not damage').

Materials and Method

Participants

The participants included 100 college students (50 females and 50 males; women's mean age: 19.72 ± 1.40 years; men's mean age: 19.74 ± 1.10 years) who had no history of neurological or psychiatric disorders based on their responses to a preliminary anamnestic questionnaire. All the participants had at least 13 years of education and provided their written informed consent before participating in the study. The study was approved by the local ethics committee of Sapienza University of Rome and was conducted in accordance with the Declaration of Helsinki.

Procedure

A psychologist presented the study to the participants and then collected informed written consent. In order to guarantee the anonymity, data collection instruments did not contain information that could identify participants.

First, participants completed the Group Embedded Figures Test (GEFT) and the self-report scales. These measures aimed at assessing cognitive style, decision-making strategies, emotion regulation and empathy.

Successively, the participants read the four moral dilemmas and answered the questions that followed each story without a time limit. The order of the stories and the protagonist's gender in the stories were counterbalanced across participants.

Measures and Materials

For the purpose of this study, we used the measures listed below, which were previously translated and validated in Italian and used in previous studies with Italian samples.

Cognitive Style Measures

We tested whether *individual cognitive style* affected the dimensions of moral reasoning using the Group Embedded Figures Test (GEFT; Witkin et al., 1971), which defines individuals as field-dependent or field-independent. The GEFT contains 18 complex figures each in which the respondent must identify a simple form.

Decision-Making Strategies

We have also assessed whether some decision-making strategies influence moral judgment. Specifically, individual preferences concerning *intuitive or deliberative decision-making* strategies were assessed with the Preference for Intuition or Deliberation Scale (PID) (Betsch and Iannello, 2010). PID conceptualizes intuition as a basic decision mode that uses affect as a decision criterion. Deliberation is defined as a decision mode following explicit evaluation, beliefs, and reasons (Betsch & Kunz, 2008). The scale comprises 18 items, 9 indicating PID-Intuition (e.g. ''My feelings play an important role in my decisions''), and 9 items indicating PID-Deliberation (e.g. ''I prefer making detailed plans rather than leaving things to chance''). Participants indicate their agreement with these statements on a 5-point Likert scale with 1 meaning "totally disagree" and 5 "totally agree". According to the method of the median split, it is possible to identify the "Intuitive" (Scoring above MDN = 30 in preference for intuition) and the "Deliberative" (Scoring below MDN = 35 in preference for deliberation) decision-making strategies.

Emotion Regulation

Individual strategies of affective regulation, such as suppression or cognitive reappraisal, were assessed with the Emotion Regulation Questionnaire (ERQ) (Gross & John, 2003). It is a 10-item scale designed to measure respondents' tendency to regulate their emotions in two ways: (1) cognitive reappraisal and (2) expressive suppression. Respondents answer each item on a 7-point Likert-type scale ranging from 1 (strongly disagree) to 7 (strongly agree).

Empathy

Furthermore, *individual differences in empathy,* that is, the reactions of one individual to the observed experiences of another (Davis, 1983), were investigated with the perspective-taking and empathic concern subscale of the Interpersonal Reactivity Index (IRI) (Davis, 1980). Perspective-taking is the tendency to spontaneously adopt the psychological point of view of others, while empathic concern assesses "other-oriented" feelings of sympathy and concern for unfortunate others. 28-items answered on a 5-point Likert scale ranging from "Does not describe me well" to "Describes me very well".

Moral Dilemmas

To formally investigate the role of gender differences in solving moral dilemmas, we asked participants to judge four different moral dilemmas in which a care-oriented or a justice-oriented behaviour could be violated (Boccia et al. 2014, 2017).

Each subject responded to four different dilemmas: two of them had women as transgressors of the moral action and two had men. To avoid an effect due to the story, we alternated the gender of the protagonist of the dilemmas. For example, for half of the respondents, a dilemma had a female protagonist named "Maria". For the other half of the respondents, the same dilemma had a male protagonist and his name was "Mario" (Mario and Maria are typical Italian names). Each moral dilemma posed a moral choice and a conflict between two possible tendencies that participants had to evaluate. The *first moral dilemma* proposed a moral choice between a crime and protection of a child, directly assessing moral attitudes towards the deontological ethic (not stealing) versus a utilitarian perspective (protecting a child using force). In the *second moral dilemma*, the participants had to judge a moral decision involving a behaviour oriented towards the utilitarian ethic (to guarantee self and community safety) versus the ethic of care (to protect the child). In the *third moral dilemma,* the participants had to judge between a care-oriented behaviour (to protect a child who stole food) and a utilitarian ethic (to identify the thief to avoid self and community punishment). In the *fourth moral dilemma,* the participants had to judge between a care-oriented behaviour (refusing to sacrifice one's child) and a utilitarian perspective (sacrificing one's child to save the other) (see *supplementary materials* for more details about the stories).

Each dilemma was followed by five questions that evaluated five aspects of the moral choice and always involved the two different orientations: (a) moral *acceptability* of the proposed choice, (b) *guilt*, (c) *responsibility* and (d) *consequences*. The participants had to express the extent to which they agreed with the question using a Likert scale that ranged from 0 to 4 where 0 indicated 'completely disagree with' and 4 indicated 'completely agree with'. Then, engagement with each moral dilemma was measured using a three-item scale. The participants were asked to recall when they were reading the dilemma and indicate to what extent they were: (1) involved, (2) interested, and (3) motivated (1 = not involved at all, not interested at all, not motivated at all; 7 = totally involved, totally interested, totally motivated) (for a similar measure, see Lee et al., 2010; Lee & Aaker, 2004). A composite score for engagement was computed by averaging the responses across the 3 items. Questions about emotional state followed each moral dilemma. The participants were asked to express how angry, sad, happy, disgusted and fearful they felt on a Likert scale ranging from 1 to 5, with 1 indicating 'not at all' and 5 indicating 'very'.

Statistical Analysis

Statistical analyses were performed using SPSS (IBM SPSS Statistics 20). We carried out one-way ANOVAs on the scores of (a) moral *acceptability* of the proposed choice, (b) *guilt*, (c) *responsibility* (d) *consequences*, (e) *engagement* and (f) *emotional state (anger, sadness*, *happiness*, *disgust and fear)* for each dilemma, with participants' gender (female and male) as the independent variable.

Then, we carried out one-way ANOVAs to determine whether scores on the PID, ERQ, IRI and GEFT showed significant differences due to participant gender. Finally, we used linear regression analyses (blocks method) to assess whether the scores on the PID, ERQ, IRI and GEFT significantly predicted the different dimensions of moral choices in the four moral dilemmas. The statistical tests we have performed are based on a number of assumptions of the data. The normal distribution, homogeneity of variances, linear relation and independence were not violated.

Results

First Moral Dilemma: Utilitarian Ethic Versus Deontological Ethic

We found a main effect of gender on moral *acceptability* ($F_{1,98} = 4.08$, $p = 0.046$), with men being more prone to accept violation of the ethic of justice (the theft) to save the child (Table 1). No effect of gender was found regarding guilt, responsibility or consequences (Table 1). We found a significant effect of gender on *engagement* ($F_{1,98} = 6.63$, $p = 0.05$), with women being more involved in the dilemma. We found a significant effect of gender on

sadness ($F_{1,97} = 8.13$, $p = 0.01$; Fig. 1A), with women being more prone to sadness. No effect of gender was found for anger, happiness, disgust or fear (Fig. 1A).

Second Moral Dilemma: Utilitarian Ethic Versus Ethic of Care

We found a main effect of gender on moral *acceptability* ($F_{1,97} = 5.35$, $p = 0.02$), with men being more prone to accept violation of the ethic of care (to kill the baby) to save themselves and others. Men also scored higher on the *consequences* scale ($F_{1,97} = 5.34$, $p = 0.02$). No effect of gender was found for guilt and responsibility (Table 1) or for engagement. We found a significant effect of gender on *sadness* ($F_{1,98} = 6.82$, $p = 0.05$) and *fear* ($F_{1,97} = 5.94$, $p = 0.05$; Fig. 1B), with women being more prone to both. No effect of gender was found for anger, happiness or disgust (Fig. 1B).

Third Moral Dilemma: Utilitarian Ethic Versus Ethic of Care

We found no main effect of gender on moral *acceptability,* guilt, responsibility or consequences in the third moral dilemma (Table 1).

We found a significant effect of gender on *engagement* ($F_{1,95} = 5$, $p = 0.05$), with women being more involved in the dilemma. We found a significant effect of gender on sadness ($F_{1,96} = 9.64$, $p = 0.00$; Fig. 1A) and *fear* ($F_{1,96} = 4.23$, $p = 0.05$; Fig. 1C), with women being more prone to sadness and fear. No effect of gender was found for anger, happiness or disgust (Fig. 1C).

Fourth Dilemma: Utilitarian Ethic Versus Ethic of Care

We found a main effect of gender on moral *acceptability* ($F_{1,95} = 5.69$, $p = 0.02$), with men being more prone to accept violation of the ethic of care (to sacrifice one child to save the other) (Table 1). We found no main effect of gender on guilt, responsibility, or consequences, but gender showed an interesting trend in the last factor, with men being more prone to consider the consequences, that is, to save the child that had the best chance of survival (Table 1).

We found a significant effect of gender on *engagement* ($F_{1,90} = 7.82$, $p = 0.01$), with women being more involved in the dilemma. We found a significant effect of gender on sadness ($F_{1,92} = 5.13$, $p = 0.05$; Fig. 1D), with women being more prone to sadness. No effect of gender was found for anger, happiness, disgust or fear (Fig. 1D).

Gender Differences in Cognitive Style, Decision-Making, Emotion Regulation and Empathy

We found no effect of gender on *cognitive style*, measured with the GEFT, or on individual preferences towards an *intuitive or deliberative decision-making* strategy, measured with the PID. Regarding the *individual strategies of affective regulation*, measured with the ERQ, we found a significant effect of gender on *suppression* ($F_{1,98} = 5.88$, $p = 0.02$) but not on cognitive reappraisal: men (mean = 16.20; S.D. = 4.96) obtained higher scores on the suppression subscale than women (mean = 13.70; S.D. = 5.35). Furthermore, we found a gender effect on *empathy* measured with the IRI: women obtained higher scores on the *empathic concern* ($F_{1,98} = 4.68$, $p = 0.03$; mean = 23.76; S.D. = 3.41) and *perspective-taking* ($F_{1,98} = 9.85$, $p = 0.00$; mean = 23.84; S.D. = 3.94) subscales (Men: empathic concern, mean = 22.32; S.D. = 3.25; perspective-taking, mean = 21.76; S.D. = 2.54).

Effect of Individual Differences in Decision-Making, Emotion Regulation and Empathy on Solving Moral Dilemmas

The participants' scores on the PID significantly predicted their scores on the consequences scale of the first moral dilemma, where a justice-oriented attitude was violated. In particular, higher scores on the *deliberative subscales* predicted higher scores on the consequences scale of the first moral dilemma (Beta = 0.26; $t = 2.76$; $p = 0.007$), and higher scores on the intuitive subscale predicted lower scores on the same consequences scale (Beta = − 0.39; $t = -4.18$; $p = 0.000$). Furthermore, higher scores on suppression, measured with the ERQ, significantly predicted higher scores on the consequences scale of the first moral dilemma (Beta = 0.303; $t = 3.17$; $p = 0.002$). In addition, higher scores on suppression also predicted higher scores on the consequences scale of the third moral dilemma (Beta = 0.22; $t = 2.24$; $p = 0.03$). We observed a significant effect of perspective-taking on the moral acceptability of the first moral dilemma (Beta = 0.23; $t = 2.20$; $p = 0.030$) and the consequences scale of the same dilemma, with a negative trend (Beta = − 0.27; $t = -2.61$; $p = 0.010$). Thus, the participants with higher scores on perspective-taking considered the consequences of the violation of a justice-oriented attitude less frequently. Empathic concern also significantly predicted the moral acceptability of the second dilemma (Beta = − 0.39; $t = -3.87$; $p = 0.000$), where a care-oriented attitude is violated, with a greater extent of rejection of the violation predicting higher scores on empathic concern. Furthermore, the participants with higher perspective-taking

Table 1 Mean and standard deviation on different dimensions of moral judgment

	Females	Males	*P*
1st moral dilemma			
Acceptability	2.52 (1.02)	2.88(0.75)	**0.046**
Guilt	3.02 (0.85)	3.12 (0.80)	0.545
Responsibility	1.60 (1.16)	1.54 (1.33)	0.810
Consequences	2.32 (1.19)	2.52 (0.97)	0.359
2nd moral dilemma			
Acceptability	1.08 (1.07)	1.59 (1.14)	**0.023**
Guilt	3.40 (0.90)	3.42 (0.64)	0.899
Responsibility	3.46 (0.91)	3.52 (0.58)	0.695
Consequences	2.14 (0.94)	2.60 (1.03)	**0.023**
3th moral dilemma			
Acceptability	1.58 (1.09)	1.54 (0.89)	0.841
Guilt	3.08 (1.03)	2.96 (0.81)	0.517
Responsibility	2.92 (1.01)	3.08 (0.80)	0.382
Consequences	1.90 (0.95)	1.86 (0.97)	0.836
4th moral dilemma			
Acceptability	1.49 (1.21)	2.13 (1.41)	**0.019**
Guilt	3.50 (0.81)	3.46 (0.78)	0.790
Responsibility	3.52 (0.61)	3.33 (0.85)	0.199
Consequences	1.94 (1.24)	2.42 (1.16)	0.053

scores felt guiltier (Beta = 0.26; $t = 2.45$; $p = 0.016$) and more responsible (Beta = 0.24; $t = 2.25$; $p = 0.027$) in the second moral dilemma, where higher empathic concern also significantly predicted lower scores on the consequences scale (Beta = − 0.30; $t = -2.87$; $p = 0.005$). Higher empathic concern scores significantly predicted lower acceptability scores (Beta = − 0.22; $t = -2.05$; $p = 0.043$) and lower scores on the consequences scale of the forth moral dilemma (Beta = − 0.22; $t = -2.05$; $p = 0.043$), where a care-oriented attitude is violated, and higher scores on perspective-taking predicted higher guilt (Beta = 0.23; $t = 2.11$; $p = 0.037$). Finally, field-independent individuals (higher scores on the GEFT) felt guiltier when a care-oriented attitude was violated (B = 0.22; $t = 2.23$; $p = 0.028$).

Discussion

Moral reasoning orientation is an important lens because it indicates a worldview that frames thinking about moral conflicts, the factors that deserve priority, and how to resolve these conflicts. This study raises some important issues about gender differences in moral choices and suggests some important implications in different fields of life, such as in economics choices or in medical decision-making.

Various interesting results emerged from the present research. The most interesting result (which confirms our hypothesis) is that the moral judgements of men and women were significantly different. For example, women judged the action of killing as less acceptable, even though it could lead to their salvation and that of others (dilemma n. 2). In contrast, men were more consequentialist in their moral choices, giving more consideration to saving themselves and the group to which they belonged. Moreover, the women considered it less morally acceptable to choose which of the two children to save when they faced an obligatory choice (dilemma n. 4), whereas the men seemed to pay more attention to the consequences (if you do not choose, both will perish). In both cases, the men's choices were more oriented towards the most beneficial utilitarian purposes, even though they might lead to violent and painful actions (to kill a little girl by smothering her or to choose which child to sacrifice). These results agreed with those of Fumagalli et al. (2010). These authors found that women were more oriented towards non-utilitarian responses in evaluating personal moral dilemmas than men.

Even in the first dilemma, the women considered the transgression less acceptable (robbing the pharmacy to obtain the medicine necessary for the child's survival, loosely based on 'The Heinz Dilemma'). This result may seem surprising and even counterintuitive. However, two aspects must be considered. First, in this case, the

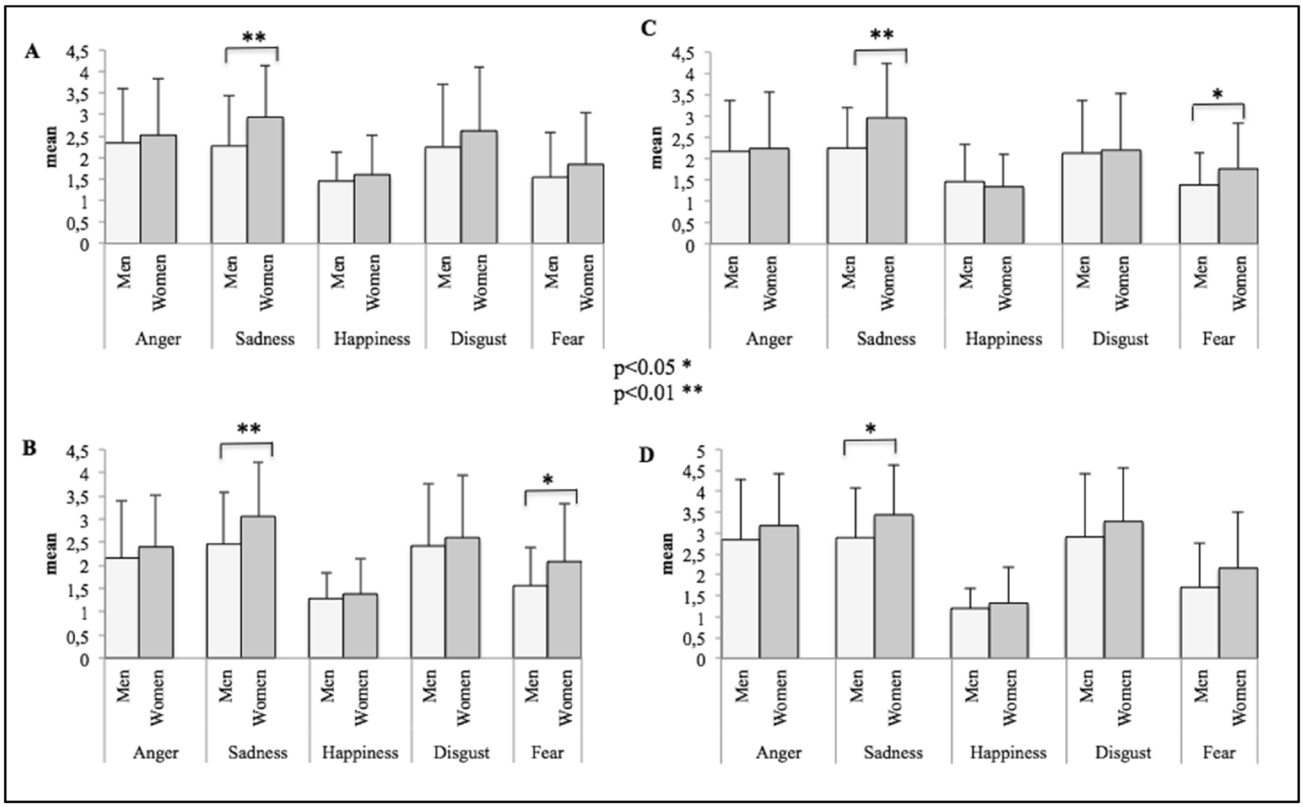

Fig. 1 Mean scores of the emotional state by gender in different dilemmas (**a**: first dilemma; **b**: second dilemma; **c**: third dilemma; **d**: fourth dilemma)

transgression involved violent behaviour (robbing the pharmacy). It must be noted that women have been found to show a lower propensity than men towards behaviours involving physical violence and are more opposed to risk (Weber et al., 2002). Second, the concept of care orientation is not exclusively related to the dyadic mother–child relationship but must be understood as a general predisposition to consider responsibility and attachment to others, with greater involvement of affective processes, such as empathy (Noddings, 2003). Our research clearly shows that men and women significantly differ in emotional involvement. The measure of engagement (involvement, motivation and interest in the story) was significantly higher in women than men (dilemmas 1, 3 and 4). Our results may be connected to gender differences in empathic ability (Baron-Cohen & Wheelwright, 2004), which make women more resistant to making decisions that are rationally viable but involve directly inflicting physical or moral pain on other individuals. Furthermore, when the women assessed their emotional state after each dilemma, they showed a greater propensity to experience emotions such as sadness (significant difference in all four dilemmas) and fear (significant in the second and third dilemmas but not in the fourth, $p = 0.07$). Therefore, women tend to experience dysphoric emotions more than men, which could affect their ethical choices. Indeed, we can interpret our results in light of the dual-process theory (Greene et al., 2001; Greene, 2009). According to Greene and colleagues, utilitarian moral choices, that is, 'utilitarian moral judgements' (or 'consequentialist' judgements) are associated with greater control of cognitive processes, whereas intuitive emotional processing is more associated with deontological judgements aimed at respecting duties and obligations.

The present results concerning gender-related differences in moral judgements are also supported by neuroimaging findings. Recent studies (Harenski & Hamann, 2006; Robertson et al., 2007; Harenski et al., 2008) have shown greater activation of brain regions in women that are generally associated with the processing of care-based information (anterior and posterior cingulate cortex, anterior insula) and increased activation of brain regions in men that are related to a justice-based orientation (posterior superior temporal sulcus).

Overall, our study supports the existence of gender differences in moral behaviour. These differences support Gilligan's theory of a 'dual moral voice' (Gilligan, 1982), but with an important clarification. The concept of care should be understood more extensively, as a predisposition to consider sensitivity and responsibility for others, with great involvement of emotional processes. In all our

dilemmas, women's choices have been more influenced by social emotions such as empathy, guilt, and shame. Care-oriented morality, driven by emotional influences, is closer to deontological ethics (Thou shall not murder! Thou shall not steal!).

Conversely, men's moral judgment is more detached from emotional processes, consequently more oriented towards cognitive processes and utilitarian choices.

Conclusion, Limits and Future Directions

We are aware that in our research there is an important criticality. This study used hypothetical dilemmas to elicit social-moral reasoning from respondents. The limitations of using hypothetical dilemmas are many. In particular is the limitation of low external validity, that refers to how well the results of a given study generalize and explain a range of other situations (Bauman et al., 2014). To facing this criticality we used two strategies. First, we have used plausible stories with good ecological validity (No fat men to push off a bridge). Second, we asked respondents how realistic the scenario was (e.g. "How realistic do you think the story is?"). The answers "the scenario is unrealistic" has eliminated. In conclusion, this study sheds more light on gender-related differences in moral reasoning. Knowing *how* individuals make their moral choices and *which* factors regulate these choices is crucial in daily life. New research should further verify whether variables such as age or cultural differences could lead to different results from those found here. Moreover, it might also be interesting to check out another important aspect. In our dilemmas, the characters of the stories acted physically violent. In the future, it could investigate whether the same gender differences are found for other types of violence, such as verbal or psychological violence.

Authors' Contributions All authors contributed equally to the work.

Funding The authors received no specific funding for this work.

Data Availability Material Generated statement: The datasets generated for this study are available on request to the corresponding author, and Material are included in the manuscript/supplementary files.

Compliance with Ethical Standards

Conflict of interest The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Human and animal Rights Statement *Studies involving animal subjects.* Generated statement: No animal studies are presented in this manuscript. *Studies involving human subjects.* Generated statement: The study was approved by the Ethics Review Board of the Department of Psychology, "La Sapienza" University of Rome. *Inclusion of identifiable human data.* Generated statement: No potentially identifiable human images or data is presented in this study.

Consent to participate Participants were informed of the aims and purpose of the study, as well as their participation rights (e.g. confidentiality of responses, allowance to leave the study at any point without any consequences), in advance of data collection. Thus, written informed consent was obtained by all the participants. None of the participants had serious accidents with traumatic consequences.

Consent for Publication All authors who have contributed to the submitted manuscript confirm that they have read and agreed to the conditions of this submission statement, including that:

- They agree to publication of this manuscript;
- the submission is original and has not been published previously;
- all permissions have been obtained;
- the manuscript includes all the relevant statements and acknowledgements.

Code Availability The authors declare that for the analysis they used IBM SPSS Statistics software, licensed by the Department of Psychology of the Sapienza University of Rome.

References

Agerström, J., Björklund, F., & Carlsson, R. (2011). Gender differences in implicit moral orientation associations: The justice and care debate revisited. *Current Research in Social Psychology, 17,* 10–18.

Aldrich, D. P., & Kage, R. (2003). Mars and Venus at twilight: A critical investigation of moralism, age effects, and sex differences. *Political Psychology, 24,* 23–40.

Aleman, A., & Swart, M. (2008). Sex differences in neural activation to facial expressions denoting contempt and disgust. *PLoS ONE, 3*(11), e3622. https://doi.org/10.1371/journal.pone.0003622.

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: An investigation of adults with Asperger syndrome or high functioning autism and normal sex differences. *Journal of Autism Developmental Disorders, 34,* 163–175.

Bauman, C. W., McGraw, A. P., Bartels, D. M., & Warren, C. (2014). Revisiting external validity: Concerns about trolley problems and other sacrificial dilemmas in moral psychology. *Social and Personality Psychology Compass, 8*(9), 536–554. https://doi.org/10.1111/spc3.12131.

Beal, C. L., Garrod, A., Ruben, K., & Stewart, T. L. (1997). Children's moral orientations: Does the gender of dilemma character make a difference? *Journal of Moral Education, 26,* 45–58.

Betsch, C., & Iannello, P. (2010). Measuring individual differences in intuitive and deliberate decision making styles: A comparison of different measures. In A. Glockner & C. Witteman (Eds.), *Tracing intuition: recent methods in measuring intuitive and deliberate processes in decision making* (pp. 251–267). London: Psychology Press.

Betsch, C., & Kunz, J. J. (2008). Individual strategy preferences and decisional fit. *Journal of Behavioural Decision Making, 21,* 532–555. https://doi.org/10.1002/bdm.600.

Bjorklund, F. (2003). Differences in the justification of choices in moral dilemmas: Effects of gender time pressure and dilemma seriousness. *Scandinavian Journal of Psychology, 44,* 459–466.

Boccia, M., Cordellieri, P., Piccardi, L., Ferlazzo, F., Guariglia, C., & Giannini, A. M. (2014). Gender differences in solving moral dilemma: Are there any implications for experts in aerospace flight? *Italian Journal of Aerospace Medicine, 10,* 12–23. https://doi.org/10.3758/s13414-014-0739-7.

Boccia, N., Verde, P., Angelino, G., Carrozzo, P., Vecchi, D., Piccardi, L., et al. (2017a). Effect of professional expertise and exposure to everyday life decision-making on moral choices. *Neuroscience Letters*, *654*, 80–85.

Boccia, M., Dacquino, C., Piccardi, L., Cordellieri, P., Guariglia, C., Ferlazzo, F., et al. (2017b). Neural foundation of human moral reasoning: an ALE meta-analysis about the role of personal perspective. *Brain Imaging and Behavior*, *11*(1), 278–292.

Botvinik, M., Jha, A. P., Bylsma, L. M., Fabian, S. A., Solomon, P. E., & Prkachin, K. M. (2005). Viewing facial expression of pain engages cortical areas involved in the direct experience of pain. *Neuroimage, 25,* 312–319.

Brabeck, M., & Shore, E. (2003). Gender differences in intellectual and moral development? The evidence that refutes the claim. In J. Demick & C. Andreoletti (Eds.), *Handbook of adult development* (pp. 351–368). New York: Kluwer Academic/Plenum Publishers.

Colby, A., & Kohlberg, L. (1987). *The measurement of moral judgment* (Vol. 1). New York: Cambridge University Press.

Davis, Mark H. (1980). A multidimensional approach to individual differences in empathy. *Journal Supplemental Abstract Service Catalog of Selected Documents in Psychology, 10,* 85.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44,* 113–126.

Friedman, W. J., Robinson, A. B., & Friedman, B. L. (1981). Sex differences in moral judgments? A test of Gilligan's theory. *Psychology of Women Quarterly, 11,* 37–46.

Fumagalli, M., Ferrucci, R., Mameli, F., Marceglia, S., Mrakic-Sposta, S., Zago, S., et al. (2010). Gender-related differences in moral judgements. *Cognitive Process, 11,* 219–226.

Garmon, L. C., Basinger, K. S., Gregg, V. R., & Gibbs, J. C. (1996). Gender differences in stage and expression of moral judgment. *Merrill-Palmer Quarterly, 42*(3), 418–437.

Gibbs, J. C., Arnold, K. D., & Burkhart, J. E. (1984). Sex differences in the expression of moral judgment. *Child Development, 55,* 1040–1043.

Gilligan, C. (1982). *In a different voice: psychological theory and women's development*. Cambridge: Harvard University Press.

Gilligan, C., & Attanucci, J. (1988). Two moral orientations: Gelnder differences and similarities. *Merrill-Palmer Quarterly, 34*(3), 223–237.

Greene, J. D. (2009). Dual-process morality and the personal/impersonal distinction: A reply to McGuire, Langdon, Coltheart and Mackenzie. *Journal of Experimental Social Psychology, 45,* 581–584.

Greene, J., & Haidt, J. (2002). How (and where) does moral judgement work? *Trend in Cognitive Sciences, 6,* 517–523.

Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron, 44,* 389–400.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). And fMRI investigation of emotional engagement. *Moral Judgment Science, 293,* 2105–2108.

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology, 85*(2), 340–362.

Gump, L. S., Baker, R. C., & Roll, S. (2000). Cultural and gender differences in moral judgment: A study of Mexican Americans and Anglo-Americans Hispanic. *Journal of Behavioral Sciences, 22,* 78–93.

Hareli, S., Shomrat, N., & Hess, U. (2009). Emotional versus neutral expressions and perceptions of social dominance and submissiveness. *Emotion, 9,* 378–384.

Harenski, C. L., Antonenko, O., Shane, M. S., & Kiehl, K. A. (2008). Gender differences in neural mechanism underlying moral sensitivity. *Social Cognitive and Affective Neuroscience, 3*(4), 13–321.

Harenski, C. L., & Hamann, S. (2006). Neural correlates of regulating negative emotions related to moral violations. *Neuroimage, 30,* 313–324.

Indick, W., Kim, J., Oelberg, B., & Semino, L. (2000). Gender differences in moral judgment: Is non-consequential reasoning a factor? *Current Research in Social Psychology, 5*(20), 1–13.

Jaffee, S., & Hyde, J. S. (2000). Gender differences in moral orientation: A meta-analysis. *Psychological Bulletin, 126,* 703–726.

Kohlberg, L. (1964a). Development of moral character and moral ideology. In M. L. Hoffman & L. W. Hoffman (Eds.), *Review of child development research* (pp. 383–430). New York: Russell Sage Foundation.

Kohlberg, L. (1964b). Development of moral character and moral ideology. In M. L. Hoffman & L. W. Hoffman (Eds.), *Review of child development research* (pp. 381–431). New York: Russell Sage Foundation.

Lee, A. Y., & Aaker, J. L. (2004). Bringing the frame into focus: The influence of regulatory fit on processing fluency and persuasion. *Journal of Personality and Social Psychology, 86,* 205–218.

Lee, A. Y., Keller, P., & Sternthal, B. (2010). Value from regulatory construal fit: The persuasive impact of fit between consumer goals and message concreteness. *Journal of Consumer Research, 36,* 735–747.

McGuire, J., Langdon, R., Coltheart, M., & Mackenzie, C. (2009). A reanalysis of the personal/impersonal distinction in moral psychology. *Journal of Experimental Social Psychology, 45,* 577–580.

Mikhail, J. (2007). Universal moral grammar: Theory, evidence and the future. *Trend in Cognitive Sciences, 11,* 143–152.

Moll, J. R., de Oliveira-Souza, R., Krueger, F., & Grafman, J. (2005). The neural basis of human moral cognition. *Nature Reviews Neuroscience, 6,* 799–809.

Noddings, N. (2003). *Caring. A feminist approach to ethics and moral education*. California: University of California Press.

Robertson, D., Snarey, J., & Ousley, O. (2007). The neural processing of moral sensitivity to issues of justice and care. *Neuropsychologia, 45,* 755–766.

Singer, T., Seymour, B., O'Doherty, J. P., Stephan, K. E., Dolan, R. J., & Frirth, C. D. (2006). Empathic neural responses are modulated by the perceived fairness of others. *Nature, 439,* 466–469.

Verde, P., Piccardi, L., Bianchini, F., Trivelloni, P., Guariglia, C., & Tomao, E. (2013). Gender effects on mental rotation in pilots vs nonpilots. *Aviation, Space and Environmental Medicine, 84,* 726–729.

Walker, L. J., de Vries, B., & Trevethan, S. D. (1987). Moral stages and moral orientations in real-life and hypotethical dilemmas. *Child Development, 58,* 842–858.

Wark, G. R., & Krebs, D. L. (1966). Gender and dilemma differences in real-life moral judgment. *Developmental Psychology, 32,* 220–230.

Weber, E. U., Blais, A. R., & Betz, N. E. (2002). A domain-specific risk-attitude scale: Measuring risk perceptions and risk behaviors. *Journal of Behavioral Decision Making, 15,* 263–290.

Witkin, H. A., Oltman, P. K., Raskin, E., & Karp, S. A. (1971). *Group embedded figures test manual*. Palo Alto: Counsulting Psychologist Press.

Yacker, N., & Weinberg, S. L. (1990). Care and justice moral orientations: A scale for assessment. *Journal of Personality Assessment, 55,* 18–27.