

Submitted : March 13, 2011  
First Revision : June 14, 2011  
Accepted : June 16, 2011

### **GENDER DIFFERENCES IN IMPLICIT MORAL ORIENTATION ASSOCIATIONS: THE JUSTICE AND CARE DEBATE REVISITED**

Jens Agerström  
Kristianstad University

Fredrik Björklund  
Lund University

Rickard Carlsson  
Lund University

#### **ABSTRACT**

*Employing new measures (Implicit Association Test) to study the classic issue of moral orientations, we predicted and found gender differences in implicit associations to the concepts of justice and care. Specifically, we found that men more strongly associate justice vs. care with importance and with themselves than women. However, participants' explicit ratings did not reveal any clear patterns of gender differences, which is consistent with previous studies. Implications for social psychological theory and research on morality are discussed.*

#### **INTRODUCTION**

The debate on gender differences in moral orientation was initiated by Carol Gilligan (1982) who claimed that as a result of being exposed to different socialization practices, women tend to construe and resolve moral dilemmas as reflecting issues of care, whereas men are more likely to construe moral issues in terms of abstract principles related to justice. Gilligan argues this is due to boys' growth being primarily concerned with separation and individuation and girls' growth being primarily concerned with attachment and relationship. As a result, men come to view themselves as independent agents, regulated only by rights and duties, whereas women come to perceive themselves as part of a network of social relations, in which the nodes are mutually responsible for one another, putting more focus on care than justice. Although this type of gender difference in moral orientation initially received empirical support, later research either failed to show it or only found weak evidence (see Jaffee & Hyde, 2000, for a meta-analysis, and Walker, 2006 for a review).

We argue that the failure to find reliable gender differences in justice and care may have to do with researchers having relied solely on measures of moral orientation that tap deliberate, reflective cognitive processes. Participants have typically been asked to self-report moral orientation or to provide resolutions to ready-made moral dilemmas after having engaged in extensive moral reasoning (e.g., Wark & Krebs, 1996; Ryan, David, & Reynolds, 2004). However, many everyday situations are likely to elicit more spontaneous moral responding where moral reasoning and reflective cognitive processes play no, or at best, a minimal role (Haidt, 2001).

### **Dual modes of morality**

Dual system models of cognition (see e.g., Chaiken & Trope, 1999, for an overview) distinguish between deliberate and automatic processes. Deliberate processes stem from rule-like learning and are based on propositional reasoning that is generally concerned with the validation of evaluations and beliefs. The subjective validity of a given proposition dictates whether it will be endorsed on an explicit level (e.g., Gawronski & Bodenhausen, 2006). Since extant measures of justice and care orientations capture these propositional processes, they should be seen as explicit measures. In contrast, automatic processes are based on non-deliberate associative activation of cognitive structures. Presumably, they evolve from associative learning as well as repeated rule based learning that over time becomes integrated with associative structures in memory. Measures such as the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) which capture people's associative structures are generally labeled implicit measures. When people declare their moral preferences or provide resolutions to justice vs. care dilemmas (explicit measures) propositional processes are likely to be employed given that such behaviors involve some deliberation. In contrast, when they engage in more spontaneous and less controlled behaviors, such as when acting altruistically on the spur of the moment, associative processes may be more influential (Perugini & Leone, 2009). In the present research, we focus on and measure gender differences in such associative processes.

Adapting Perugini and Leone's (2009) implicit moral concept model to the justice and care domain, we believe that the implicit self-concept label ("Just" and "Caring") evolves from associations between actions (reciprocate, empathize) and the implied trait-concept (Just, Caring) that the individual accrues throughout a history of repeated social interactions. According to this view, the implicit self-concept gradually evolves as repeated associative learning processes take place. If a person increasingly engages in actions that are associated with Justice or Care, his or her implicit self-concept will be increasingly polarized as Just or Caring, respectively. The model holds that associations among self, moral orientation, and action schemas depicting actions pertinent to the specific self-concept, are responsible for the behavioral consequences of this associative network. The strength of the self-behavior associations in the person's mind is contingent on frequent co-activation of the action (reciprocate vs. helping the needy) and the trait-based concept label underlying the behavior (Just vs. Caring). Importantly, these associations may also develop from a socialization process where norms (e.g., prescriptive stereotypes of how women and men should behave) in society help shape these associations. Regardless of the reasons, women and men may come to have different implicit associations of justice and care.

## The Present Research

In the present study we approach the justice versus care debate from a new angle by examining whether gender differences in moral orientation exist on the automatic associative level. More specifically, using a “moral value” version of the IAT, we predicted that females would more easily associate care concepts with attributes denoting significance and justice concept with insignificance relative to males who would more easily associate justice with significance and care concepts with insignificance. Moreover, using a “Self” vs. “Other” self-concept IAT, it was predicted that women would more easily associate care concepts with themselves and justice concept with others relative to men who should more easily associate justice with themselves and care concepts with others. To our knowledge this is the first study to examine gender differences in moral orientation using implicit measures. Hopefully, this research will shed new light on and spur more studies on a classic research topic within the psychology of morality that has come to a halt.

## STUDY 1

### METHOD

#### *Participants and Procedure*

Seventy-four students (31 males, 43 females; mean age 22.36 years,  $SD = 2.59$ ) were approached on the social sciences’ campus and asked to participate in a study on moral perception. They first completed the self-concept IAT followed by the value-IAT and lastly filled out the explicit measures and provided demographic data. Afterwards they were debriefed and thanked for their participation.

#### *Measures and Materials*

The IAT (Greenwald et al., 1998) is a computerized response latency measure designed to tap individual differences in automatic associations between categories (e.g., flowers vs. insects) and attributes (e.g., good vs. bad). The IAT requires the rapid classification of various stimuli and rests on the logic that easier pairings (i.e., faster responses) reflects stronger associations compared to difficult pairings (i.e., slower responses). In the present implementation of the IAT, the category labels were always “Justice” versus “Care”. The stimuli words reflecting “Justice” were *principles, fairness, honesty, and rules*, whereas words reflecting “Care” were *empathy, unselfishness, consideration, and solidarity*. The value-IAT had “important” vs. “unimportant” as attribute labels with the corresponding stimuli *central, significant, relevant, and essential* vs. *secondary, insignificant, irrelevant, and unessential*. The self-concept IAT used the attribute labels “Self” vs. “Other”, with *I, me, myself* and *they, you, them* as stimuli. An overview of the IAT procedure can be found in the Appendix.

The IAT effects were calculated (using the *D*-scoring algorithm recommended by Greenwald, Nosek, and Banaji, 2003) so that a positive *D*-value reflected a stronger association of justice than care with importance (value-IAT) or the self (self-concept IAT). The reliabilities of the IATs were acceptable;  $\alpha = .74$  for the self-concept IAT and  $\alpha = .78$  for the value-IAT.

The explicit measures were constructed to be conceptually similar to the IATs. As for the explicit self-concept, participants indicated on a 6-point scale ranging from 1 (*not at all*) to 6 (*very much*) to what extent they associate themselves with *principles, fairness, honesty, and rules* ( $\alpha = .71$ ), on the one hand, and *empathy, unselfishness, consideration, and solidarity* ( $\alpha = .84$ ), on the other. Then participants rated each of these justice ( $\alpha = .68$ ) and care concepts ( $\alpha = .79$ ) with respect to how important they are to them as individuals on a 6-point scale ranging from 1 (*not at all important*) to 6 (*extremely important*).

## RESULTS

### *Implicit Measures*

To test the hypothesis that women automatically associate themselves with care (relative to justice) to a greater extent than men, we analyzed the *D*-values of the self-concept IAT, by means of a 2 (order of compatible and incompatible part in the IAT)  $\times$  2 (gender) between-participants ANOVA. Neither the main effect of order or the interaction of order and gender were significant, ( $F(1, 70) = 2.49, p = .119$ , partial  $\eta^2 = .03$  and  $F(1, 70) = 1.71, p = .196$ , partial  $\eta^2 = .02$ , respectively), ruling out an order confound of the gender differences. As predicted, the main effect of gender was significant,  $F(1, 70) = 4.27, p = .042$ , partial  $\eta^2 = .06$ . A closer inspection of the means shown in Table 1 reveals that although both women and men automatically associate themselves with relatively more care than justice, women have a stronger relative association of care and the self than men do.

Our next hypothesis was that this pattern of relative gender differences also would show up in an value IAT that captured associations of care vs. justice with importance. Thus, we analyzed the value-IAT results (*D*-values) by means of a 2 (order of compatible and incompatible part in the IAT)  $\times$  2 (gender) between-participants ANOVA. Neither the main effect of order nor the interaction effect of order and gender was significant (Both  $F_s = 0.01$ ) and their effect sizes were virtually non-existent (both partial  $\eta^2 = .00$ ). As predicted, the main effect of gender was significant,  $F(1, 68) = 4.24, p = .043$ , partial  $\eta^2 = .06$ . Looking at Table 1, we see that the women show virtually no difference in how they automatically associate justice and care with importance whereas the men associate justice (relative to care) with higher importance.

**Table 1. Means and standard deviations of the IATs as a function of gender**

Implicit measure	Mean <i>D</i> -scores and standard deviation	
	Males	Females
Self-concept IAT	-0.14 (.38)	-0.31 (.38)
Value IAT	0.24 (.53)	-0.01 (.47)

*Note.* A positive *D*-score means a stronger association with justice relative to care, whereas as negative *D*-score represents the opposite (stronger association with care relative to justice).

The two IATs were moderately correlated,  $r(72) = .37, p = .001$ . In order to explore whether the correlations between the IATs differed across gender, we used Fisher's  $r$  to  $z$ -transformations. Interestingly, there was a trend for the correlations to be stronger for the women ( $r = .50$ ) relative to the men ( $r = .13$ ),  $z = -1.68, p = 0.084$ . Although the finding is uncertain at this point, it tentatively suggests that women have a stronger association between how they automatically associate themselves with care and how important they (automatically) find care to be.

### ***Explicit Measures***

The next step of the analysis was to investigate if the sex difference found on the implicit level also would show up in the explicit measures. The explicit importance scales were analyzed by means of a 2 within (scale: justice vs. care)  $\times$  2 between (gender) mixed ANOVA. There was a significant main effect regarding the difference between the two scales,  $F(1, 72) = 92.57, p < .001$ , partial  $\eta^2 = .56$ , suggesting that the participants overall rated care as more important than justice. Further, the main effect of gender was marginally significant,  $F(1, 72) = 2.91, p = .092$ , partial  $\eta^2 = .04$ , with females providing higher ratings overall than males. However there were no significant interaction between gender and scale,  $F(1, 72) = 1.11, p = .300$ , partial  $\eta^2 = .015$ . In other words, males did not explicitly value justice relative to care more than females, as they did implicitly.

To examine gender differences in explicitly stated self-concepts the self-ratings were analyzed by means of 2 within (scale: justice vs. care)  $\times$  2 between (gender) mixed ANOVA. There was a significant main effect regarding the difference between the two scales,  $F(1, 72) = 42.51, p < .001$ , partial  $\eta^2 = .37$ , suggesting that the participants overall associated themselves more with care than with justice. Furthermore, there was a marginally significant main effect of gender,  $F(1, 72) = 3.59, p = .062$ , partial  $\eta^2 = .05$ , indicating that women explicitly stated a stronger association with the self and moral orientations regardless of dimension. However, there was no significant interaction between gender and scale,  $F(1, 72) = .50, p = .480$ , partial  $\eta^2 = .01$ . Put differently, females did not explicitly associate themselves more strongly with care relative to justice than did males, which they did on the implicit level. Table 2 presents the means and standard deviations for the explicit measures as a function of gender.

**Table 2. Means and standard deviations of explicit measures as a function of gender**

Explicit measure	Means and standard deviation	
	Males	Females
Importance of justice	3.98 (.76)	4.12 (.83)
Self-association with justice	3.77 (.81)	3.97 (.82)
Importance of care	4.85 (.82)	5.21 (.60)
Self-association with care	4.44 (.87)	4.81 (.77)

### ***Relationship between Explicit and Implicit Measures***

Because the two IATs captured the relative associations of care vs. justice and importance (vs. unimportant) or self (vs. others), respectively, we calculated difference scores of the explicit scales. These two explicit difference scores strongly correlated with each other,  $r(74) = 0.78, p < .001$ . Regarding the self-concept IAT, it is significantly and moderately correlated with both the explicit self-concepts,  $r(74) = 0.31, p = .001$  and the explicit importance scale,  $r(74) = .35, p = .002$ . The importance-IAT is, on the other hand, not significantly correlated with the explicit importance scale,  $r(72) = .17, p = .159$ , and marginally significantly correlated with the explicit self-concept scale,  $r(72) = .23, p = .056$ .

## **STUDY 2**

Because we were somewhat concerned that the value-IAT could have been affected by the participants completing the self-concept IAT first, we wanted to replicate this IAT in a new sample. The only difference in this replication was the new sample of 80 undergraduates (34 males, 46 females; mean age 24.35 years,  $SD = 3.73$ ), who were recruited in the same way as in Study 1.

As in Study 1, the value-IAT results ( $D$ -values) were analyzed by means of a 2 (order of compatible and incompatible part in the IAT)  $\times$  2 (gender) between-participants ANOVA. Both the main effect of order and the interaction of order and gender were non-significant and had small effect sizes with  $F(1, 76) = 0.167, p = .684, \text{partial } \eta^2 = .002$  and  $F(1, 76) = 0.151, p = .699, \text{partial } \eta^2 = .002$  respectively. Replicating the finding of Study 1, the main effect of gender  $F(1, 76) = 9.837, p < .01, \text{partial } \eta^2 = .115$  was significant. The means indicate that both females and males automatically associate justice with higher importance than care, but that this association is weaker for females (mean  $D = .122, SD = 0.39$ ) than for males (mean  $D = .406, SD = 0.39$ ). Although females have a slightly more justice-oriented association in this replication so do the males by a similar amount, leaving the relative difference between the two genders virtually the same in this new sample. The most plausible explanation for this is that the effects were somewhat weaker in the first study because the participants had already completed a very similar IAT.

## GENERAL DISCUSSION

The present study revisited the classic issue of gender differences in moral orientation. Employing implicit measures, we predicted and found that men and women differ both in the extent to which they associate justice vs. care with importance (tapping their implicit evaluations of the moral orientations) and with themselves (tapping their implicit moral orientation self-construct). Specifically, we found that women had more care-oriented (relative to justice) implicit associations than the men had. Further, the importance-IAT and the self-IAT correlated moderately, a correlation which tended to be stronger for the women. Finally, participants' explicit ratings of how important justice and care is and to what extent they associate themselves with justice vs. care did not show this pattern of gender differences, but rather suggest that women gave higher ratings in general.

Our results are novel in that they appear to be the first showing gender differences in implicit associations related to justice and care. As for their theoretical implications, we observe that the relative gender difference found here are not equivalent to those suggested by Gilligan (1982), i.e. that men primarily have a moral orientation related to justice whereas women primarily have a moral orientation related to care. In fact, although men associated themselves less with care than women did, the men in the present study did not associate themselves more with justice than care. One reason for why this is so may be the sample, Sweden being a country stressing equal opportunity and treatment, including socialization practices such as upbringing and education. Hence, it should come as no surprise that men too appreciate and associate themselves with care-related moral norms and that women associate justice with high importance. However, our results suggest that the relative differences between men and women still exist.

Because the IAT is inherently a relative measure, we suggest that the absolute values of the differences in associations of justice vs. care should be interpreted with some caution. Importantly, it is impossible to tell whether the differences are mainly due to differences in associations towards care or towards the associations of justice. Although this does not change our main conclusions regarding the relative differences between men and women, it limits our ability to pin-point their exact nature. New methods related to the IAT (such as the Sorting Paired Features test) are being developed and may be more suitable for making absolute comparisons in the future, but at the present point we are reluctant to draw any conclusions beyond those related to relative effects.

As noted in the introduction, previous studies have frequently failed to find gender differences in moral orientation, which may have contributed to cooling the interest for this issue. The present findings are interesting in this context; using a measure that primarily taps automatic processes significant gender differences are revealed, suggesting implicit methods to be viable tools for further research. One explanation of the scarcity of findings related to gender differences in moral orientation may thus be the exclusive focus on measures related to deliberative processes. Besides overlooking the role of implicit cognitive processes in moral orientation which may affect moral responding in a variety of situations, it is possible that the results from extant research examining gender differences in justice and care could to some extent be plagued by

self-presentation. Women may be reluctant to explicitly associate themselves with traditional stereotype-congruent attributes and vice versa for men, and may thus adjust their moral responding so that it does not conform to traditional gender stereotypes (e.g. female = nurturing). This could mask and underestimate true gender differences in moral orientation. It may be assumed that people's implicit associations concerning gender are more resistant to change than are their explicit counterparts and thereby less affected by gender-role challenges.

To conclude, these types of measures may not only be useful in detecting gender differences in moral orientation, but also for studying individual differences. Consequently, future studies on how people construe and resolve moral problems may do well to also include these types of implicit measures.

## REFERENCES

- Chaiken, S., & Trope, Y. (Eds.). 1999. *Dual process theories in social psychology*. New York: Guilford Press.
- Gawronski, B., & Bodenhausen, G. V. 2006. "Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change". *Psychological Bulletin*, 132, 692–731.
- Gilligan, C. 1982. *In a different voice*. Cambridge, MA.: Harvard University Press.
- Greenwald, A., McGhee, D., & Schwartz, J. 1998. "Measuring individual differences in implicit cognition: The Implicit Association Test". *Journal of Personality and Social Psychology*, 74, 1464–1480.
- Greenwald, A., Nosek, B., & Banaji, M. 2003. "Understanding and using the Implicit Association Test: I. An improved scoring algorithm". *Journal of Personality and Social Psychology*, 85, 197-216.
- Haidt, J. (2001). "The emotional dog and its rational tail: A social intuitionist approach to moral judgment". *Psychological Review*, 108, 814-834.
- Jaffee, S., & Hyde, J. S. 2000. "Gender differences in moral orientation: a meta-analysis". *Psychological Bulletin*, 126, 703-726.
- Perugini, M., & Leone, L. 2009. "Implicit self-concept and moral action". *Journal of Research in Personality*, 43, 747-754.
- Ryan, M. K., David, B., & Reynolds, K. J. 2004. "Who cares? The effect of gender and context on the self and moral reasoning". *Psychology of Women Quarterly*, 28, 246-255.
- Walker, L. J. (2006). "Gender and morality". In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp. 93-115). Mahwah, NJ: Erlbaum.

Wark, G. R., & Krebs, D. L. 1996. "Gender and dilemma differences in real-life moral judgment". *Developmental Psychology*, 32, 220-230.

## AUTHOR BIOGRAPHIES

Jens Agerström is an Assistant Professor at Kristianstad University and conducts research on moral cognition, stereotypes and discrimination. E-mail: jens.agerstrom@hkr.se

Fredrik Börklund is an Associate Professor at Lund University whose research concerns moral cognition, prejudice and personality assessment. E-mail: fredrik.bjorklund@psychology.lu.se

Rickard Carlsson is a doctoral student at Lund University researching mixed stereotypes, discrimination and implicit measures pertaining to these. E-mail: [rickard.carlsson@psychology.lu.se](mailto:rickard.carlsson@psychology.lu.se)

## Appendix. Overview of the IAT procedure

Part #	Task	Category Label #1 (Typed Key = D)	Stimuli	Category Label #2 (Typed Key = K)
1	Category practice	Justice	Honesty vs. Empathy	Care
2	Attribute Practice	Important	Significant vs. Irrelevant	Unimportant
3	Critical part #1	Justice or Important	Honesty vs. Empathy or Significant vs. Irrelevant	Care or Unimportant
4	Reverse Keys practice	Care	Honesty vs. Empathy	Justice
5	Critical part #2	Care or Important	Honesty vs. Empathy or Significant vs. Irrelevant	Justice or Unimportant

*Note.* The practice parts consisted of 20 trials, whereas the critical parts each consisted of two blocks of 20 + 40 trials. Which combination (e.g., Justice + Important/ Care + Unimportant) was presented first was counterbalanced between participants and also controlled for in the analyses (referred to as the order effect). The self-concept IAT used the attribute labels "self" and "other" instead of "important" and "unimportant" but was otherwise identical. All stimulus words were carefully pre-tested to ensure that they clearly belonged only to one category label.