

Is utilitarian sacrifice becoming more morally permissible?

Ivar R. Hannikainen, Edouard Machery, & Fiery A. Cushman

A central tenet of contemporary moral psychology is that people typically reject active forms of utilitarian sacrifice. Yet, evidence for secularization and declining empathic concern in recent decades suggests the possibility of systematic change in this attitude. In the present study, we employ hypothetical dilemmas to investigate whether judgments of utilitarian sacrifice are becoming more permissive over time. In a cross-sectional design, age negatively predicted utilitarian moral judgment (Study 1). To examine whether this pattern reflected processes of maturation, we asked a panel to re-evaluate several moral dilemmas after an eight-year interval but observed no overall change (Study 2). In contrast, a more recent age-matched sample revealed greater endorsement of utilitarian sacrifice in a time-lag design (Study 3). Taken together, these results suggest that today's younger cohorts increasingly endorse a utilitarian resolution of sacrificial moral dilemmas.

Keywords: moral dilemmas, cohort effect, aging, utilitarianism.

In 1967, the philosopher Philippa Foot published an essay on an obscure ethical principle, the *doctrine of double effect*. Her essay introduced now-famous cases like the trolley problem in order to crystalize the competing mandates of deontology (never to use someone as a means to an end) versus utilitarianism (to promote the good of the many). Next, it argued that moral judgments regarding abortion and euthanasia reflect this precise tension.

In the decades since, public attitudes toward euthanasia and especially abortion have become substantially more permissive (Inglehart, 1997; Norris & Inglehart, 2011). Might these developments reflect a deeper, more systematic shift in the relative balance of deontological versus utilitarian concerns? To answer this question, we assess evidence for historical change in the way that people resolve the kinds of moral dilemmas posed by Foot fifty years ago.

Two lines of evidence motivate the prediction that utilitarian moral values are on the rise. The first concerns cohort changes in trait empathy and their predicted consequences for moral psychology. In

large-scale cross-sectional studies (total $N > 70,000$) of the Interpersonal Reactivity Index (Davis, 1980)—a multidimensional measure of self-reported affect—younger participants report lower scores on the empathic concern subscale than either middle-aged or older adults (O'Brien, Konrath, Gröhn, & Hagen, 2013). Meanwhile, a cross-temporal meta-analysis of 72 administrations of the IRI among United States college students revealed a general decline in self-reported empathy between 1979 and 2009 (Konrath, O'Brien & Hsing, 2011). This generational trend predicts a weakening prohibition on utilitarian sacrifice, since dispositional empathy—as reported on the IRI (Gleichgerrcht & Young, 2013; Patil & Silani, 2014)—is linked to deontological reactions to the trolley problem.

Second, numerous Western cultures have undergone processes of *secularization* (Norris & Inglehart, 2011), characterized by religious disaffiliation and declines in church attendance (Schwadel, 2010). In turn, studies in moral psychology reveal that religious believers are more likely to

oppose utilitarian sacrifice (Conway & Gawronski, 2013; Piazza & Landy, 2013)—a pattern which may arise from their more intuitive cognitive style (Shenhav, Rand & Greene, 2012) and a corresponding preference for the intrinsic moral evaluation of acts (Hannikainen, Miller & Cushman, 2017). Together these results provide additional grounds to suspect that utilitarian ethics may be proliferating, at least in secularizing societies.

Motivated by these existing lines of evidence, we examine the hypothesis that utilitarian moral values are spreading over time. Our methods are based on three complementary designs: In Study 1, we evaluate the relationship between age and moral judgment in a *cross-sectional* design, i.e., comparing the moral judgment of different age groups at a fixed point in time. Next, in Study 2, we examine changes in moral judgment over the human life span in a fixed panel adopting a *longitudinal* design. Finally, in Study 3, we employ a *time-lag* approach, comparing the moral judgment of similar age groups at different periods.

Consistent effects in longitudinal and cross-sectional designs (e.g., greater condemnation of utilitarian sacrifice in older age) can be treated as indicative of a *maturation* effect—i.e., that processes of aging promote deontological views. If time-lag and cross-sectional analyses reveal consistent findings, a predominant effect of *cohort* may be assumed to be present (such as greater endorsement of utilitarian sacrifice among recent generations). Finally, convergent effects in longitudinal and time-lag studies (e.g., more utilitarian judgment in recent test administrations) would suggest the prevailing influence of a *period* effect—changes resulting from the passage of time that affect everyone alike.

General Methods

All reported studies were approved by the institutional review board at Harvard University.

Participants

1. 2007-08 wave.

Between October of 2007 and June of 2008, 4,134 volunteers (1662 women, 2472 men; age IQR: 20 – 37), took part in the present study. Many participants were either college students (“*Some college*”: 1052 [25%]) or graduates (“*Bachelor’s degree*”: 968 [23%]), and almost half of all participants were US nationals (1898 [46%]). Many other participants came from Australia (119 [3%]), Canada (162 [4%]), Germany (233 [6%]), Poland (284 [7%]), and the United Kingdom (301 [7%]). Approximately half of the participants reported no religious affiliation (“*None*”: 2051 [50%]), and many others were of Christian denomination (“*-Catholic*”: 512 [12%]; “*-Orthodox*”: 55 [1%]; “*-Other*”: 329 [8%]; “*-Protestant*”: 409 [10%]).

2. Longitudinal panel.

Between July 2016 and March 2017, we re-contacted all 752 participants from the 2007-08 administration who voluntarily provided their e-mail address to take part in future research: 166 (22%) e-mails bounced, 161 (21%) participants started the survey and, after excluding 21 incomplete participations, our re-test sample consisted of 123 participants (73 men, 50 women; retest age IQR: 34 – 55) born between 1930 and 1994. Further demographic information was retrieved from the first phase: United States was the primary nationality (53 [41%]), followed by Germany (13 [10%]), United Kingdom (13 [10%]), Poland (8 [6%]) and Canada (6 [5%]). Many participants reported no religious affiliation (“*None*”: 80 [49%]), and some were Christian (“*-Catholic*”: 15 [12%]; or “*-Protestant*”: 12 [10%]). Many participants held either a Master’s (36 [28%]) or a Bachelor’s (31 [24%]) degree. Tests of attrition bias are reported in Supplementary Analysis 1.

3. 2015-17 wave.

Between October of 2015 and March of 2017, 9337 volunteers (4825 women, 4076 men, 130 other, 156 preferred not to specify; age IQR: 19 – 28) took part in the present study. Most participants were either high school graduates (“*High school/GED*”: 2018 [22%]), college students (“*Some college*”: 2481 [27%]) or graduates (“*Bachelor’s degree*”: 1956 [21%]). Over half of all participants were US nationals (5314 [57%]). Many other participants came from Australia (292 [3%]), Canada (521 [6%]), Germany (214 [2%]), and the United Kingdom (608 [7%]). More than half of the participants reported no religious affiliation (“*None*”: 5397 [58%]), and many others were of Christian denomination (“*-Catholic*”: 976 [10%]; “*-Orthodox*”: 153 [2%]; “*-Other*”: 790 [8%]; “*-Protestant*”: 707 [8%]).

Procedure

Participants visited the Moral Sense Test website (www.moralsensetest.com), either voluntarily (in Studies 1 and 3) or upon receiving an e-mail request (Study 2). After providing informed consent, participants completed at least the following three sections:

1. *Moral dilemmas.*

Participants viewed a battery of thirteen, high-conflict personal dilemmas, previously employed in numerous studies in moral psychology (see Koenigs et al., 2007). Each hypothetical situation was narrated in the second person (placing the reader in the role of actor), and presented a dilemma whether to personally sacrifice someone in order to save a larger number of lives. After each dilemma, participants were asked to rate the permissibility of the utilitarian action on a seven-point scale, anchored at (1) “*forbidden*”, (4) “*permissible*”, and (7) “*obligatory*”. We calculated a moral judgment index per participant, by averaging permissibility ratings across all thirteen dilemmas, such that higher values indicate more characteristically utilitarian moral views. The moral judgment index

revealed very good internal consistency in the present studies (Cronbach’s $\alpha \geq .87$) and also test-retest reliability in Study 2 ($r = .67$).

2. *Interpersonal Reactivity Index.*

Participants completed a widely-used assessment of self-reported empathy (Davis, 1983). The IRI contains 28 items, organized in four subscales:

- a. *perspective-taking*, the tendency to evaluate situations from the point of view of others (e.g., “I try to look at everybody’s side of a disagreement before I make a decision”);
- b. *fantasy*, the capacity to transpose oneself into the feelings and actions of characters in fictional contexts (e.g., “After seeing a play or movie, I have felt as though I were one of the characters”);
- c. *empathic concern*, the tendency to feel compassion and concern for others (e.g., “I often have tender, concerned feelings for people less fortunate than me”); and
- d. *personal distress*, own feelings of unease and discomfort in reaction to the emotions of others (e.g., “Being in a tense emotional situation scares me”).

3. *Demographic information.*

Participants were asked to provide information about: their gender; age (in years); educational attainment (1: “*Less than high school*” – 5 : “*Graduate degree*”); religious affiliation (“*Buddhist*”, “*Christian – Catholic*”, “*Christian – Orthodox*”, “*Christian – Other*”, “*Christian – Protestant*”, “*Hindu*”, “*Jewish*”, “*Muslim*”, “*None*”, “*Other*”); and religious self-identification (anchored at 1: “*Not religious*”, and 7: “*Very religious*”). Stimuli, data and scripts are available online at osf.io/ks3wz/.

Power analysis

Given our large sample sizes in Studies 1 and 3, our analyses were highly-powered to detect small effects ($r = .10$, $d = 0.20$): i.e., $\alpha < .001$ (consistent with Benjamin et al.,

2017), and $1 - \beta > .99$, setting β/α ratio to 4:1. Our longitudinal study depended on a more limited sample of 123 re-test participants. With $\alpha = .05$, and $1 - \beta = .80$, our planned analysis (paired *t*-test) afforded us enough statistical power to detect effects larger than or equal to Cohen’s $d = 0.25$.

We adopt pairwise deletion throughout this report: Each statistical analysis includes *all* participants for whom the data are available, resulting in some variation in sample sizes across analyses.

Study 1: Cross-sectional age differences

In Study 1, we examine the relationship between participants’ age and their

judgments about the permissibility of utilitarian sacrifice. If either *maturation* or *cohort* effects are present, we should observe a correlation. In contrast, if only *period* effects are present, we should observe no differences in moral judgment by age.

We also seek to replicate previously reported relationships between empathic concern and both age (O’Brien et al., 2013) and moral judgment (Gleichgerrcht & Young, 2013).

Results

Summary statistics are displayed in Table 1 below.

Table 1. 2007-08 wave: Summary statistics and correlation table.

	Mean	SD	(1)	(2)	(3)	(4)	(5)	(6)
(1) Age	30.3	13.1	--					
(2) Moral judgment	3.86	1.37	-.23	--				
(3) Fantasy	3.43	0.83	-.14	-.00 #	--			
(4) Empathic concern	3.65	0.77	.18	-.29	.33	--		
(5) Perspective-taking	3.54	0.76	.12	-.11	.21	.47	--	
(6) Personal distress	2.59	0.79	-.14	-.08	.17	.09	-.13	--
(7) Religiosity	2.70	1.87	.08	-.25	.06	.20	.06	.02 #

Note. # indicates non-significance. Otherwise $p < .005$.

Empathic concern and moral judgment.

Consistent with prior research, the condemnation of utilitarian sacrifice (Cronbach’s $\alpha = .90$) was associated with greater self-reported empathy, $r(3810) = -.29$ [-.32, -.26], $p < .001$. Controlling for the remaining IRI subscales in a multiple linear regression, the effect of empathic concern remained significant, $B = -0.58$, $t = -18.09$, $p < .001$, $\eta_p^2 = .087$.

Empathic concern and age.

We also replicated past findings concerning age differences: Age correlated positively with empathic concern, $r(3823) = .18$ [.15, .21], $p < .001$. In a polynomial model, age (centered) demonstrated both linear, $B = 0.013$, $t = 10.29$, $p < .001$, $\eta_p^2 = .031$, and quadratic, $B = -1.94 \times 10^{-4}$, $t = -3.27$, $p < .001$, $\eta_p^2 = .003$, effects on self-reported

empathy (see Figure 1B). An F-test indicated that the polynomial model ($AIC_2 = 8770$) provided better fit than the simple linear model ($AIC_1 = 8778$), $F(1, 3822) = 10.7$, $p = .001$.

Age and moral judgment.

Our primary analysis indicated that younger participants were more likely to endorse utilitarian sacrifice, $r(3808) = -.23$ [-.26, -.20], $p < .001$. Given that age revealed a quadratic effect on empathic concern, we tested a corresponding model with moral judgment as the dependent measure: Indeed, age (centered) exhibited both linear, $B = 0.03$, $t = -14.4$, $p < .001$, $\eta_p^2 = .052$, and quadratic, $B = -6.02 \times 10^{-4}$, $t = -5.77$, $p < .001$, $\eta_p^2 = .009$, effects on moral judgment (see Figure 1A). As with empathic concern, removing the quadratic term significantly

reduced model fit ($AIC_2 = 12955$, $AIC_1 = 12986$), $F(1, 3807) = 33.3$, $p < .001$.

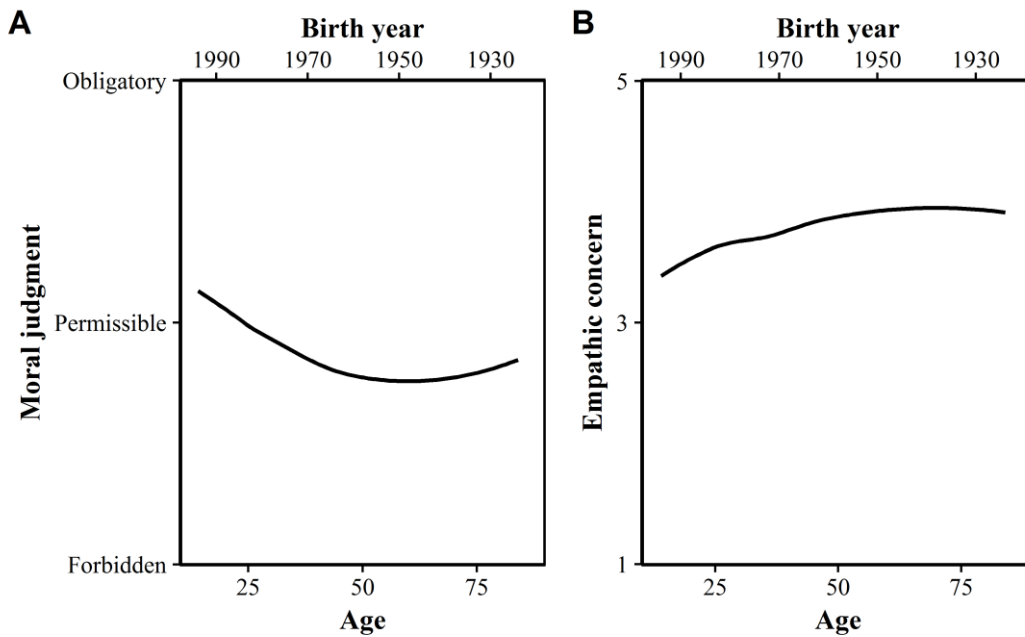


Figure 1. Age curve of (A) moral judgment and (B) empathic concern, with locally-weighted (loess) smoothing. The secondary x-axis displays corresponding birth years.

Study 2: Longitudinal analyses

Study 1 revealed differences in moral judgment as a function of participants' age: Namely, younger participants reported greater approval of utilitarian sacrifice than did older participants. As noted, this result may reflect differences between cohorts, the influence of maturation, or both.

Longitudinal designs offer a window into maturation effects, by examining patterns of change *within* individuals. So, in Study 2, we re-contacted a panel of visitors to the Moral Sense Test website and invited them to take part in an identical retest, with at least an eight-year interval between phases.

If processes of aging magnify attitudes of moral opposition toward utilitarian sacrifice, we should observe greater moral condemnation (and perhaps also empathic concern) in the second phase of this study than in the first phase—a result that would suggest that the age differences in Study 1 are driven by maturation.

Results

Manipulation check

Overall, participants had aged an average of 8.4 years between phases ($age_{2007-08} Mdn = 32$, $Q_1 = 26$, $Q_3 = 47$; $age_{2016-17} Mdn = 41$, $Q_1 = 34$, $Q_3 = 55$; $CLES = 0.68$), yielding a probability of .68 that randomly selecting an individual from each phase would return an older participant in the second phase.

Maturation effects

IRI scores (FS $\alpha = .81$; EC $\alpha = .81$, PT $\alpha = .80$, PD $\alpha = .83$) were highly correlated between phases (FS $r = .80$, EC $r = .63$, PT $r = .66$, PD $r = .62$, all $ps < .001$). No significant changes in empathic concern, $t(122) = -0.28$, $p = .78$, $d = 0.02$, or perspective-taking, $t(122) = 0.99$, $p = .33$, $d = 0.09$, were observed. However, we did find a significant decrease in fantasy, $t(122) = -3.66$, $p < .001$, $d = -0.33$, and an increase in personal distress, $t(122) = 2.94$, $p = .004$, $d = 0.26$ (see also Supplementary Table 1).

Similarly, religiosity was highly correlated between phases, $r(123) = .73$ [.65, .81], $p < .001$. A paired *t*-test revealed a significant decline, ($religiosity_{2007-08} = 2.37$, $SD = 1.78$; $religiosity_{2016-17} = 2.00$, $SD =$

1.70), $t(122) = -3.18, d = -0.29 [-0.46, -0.11], p = .002$. In Study 1, we saw a weak *positive* correlation between religiosity and age, suggesting that the present decline within individuals is attributable primarily to the influence of period (and not of maturation; see Schwadel, 2010).

Finally, moral judgment ($MJ_{2016-17} \alpha = .87$) also correlated strongly between phases, $r(123) = .67 [.56, .76], p < .001$.

Critically, a paired t -test revealed no significant shift in moral judgment, ($MJ_{2007-08} = 4.02, SD = 1.38; MJ_{2016-17} = 3.92, SD = 1.18$), $t(122) = 1.11, d = -0.10 [-0.28, 0.08], p > .250$. From a Bayesian perspective, these results provide “positive” or “substantial” support ($BF_{01} = 5.71$) for the absence of an aging effect, and the result was robust to different Cauchy prior widths (maximum $BF_{10} = 1$).

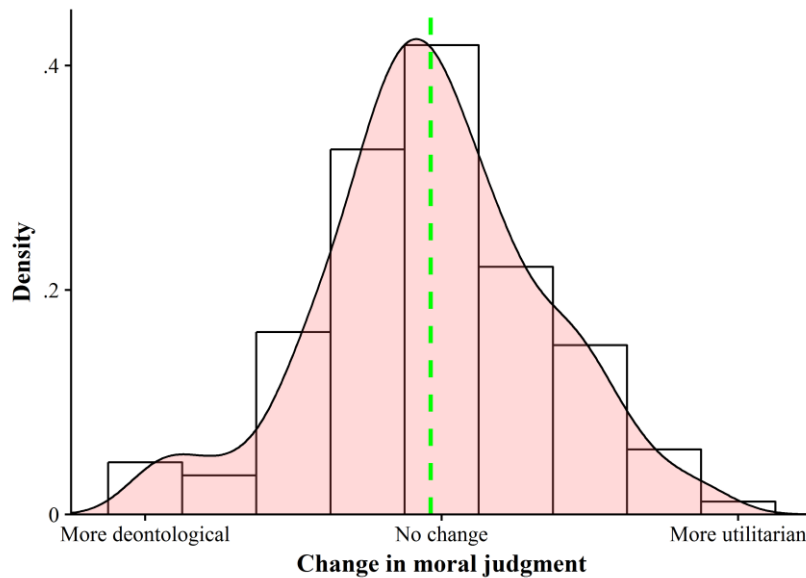


Figure 2. Histogram and density plot of longitudinal change in moral judgment. A dashed line depicts the mean level of change.

Replicating Study 1, empathic concern predicted deontological judgment in both phases (2007-08: $B = -0.69, t = -4.00, p < .001, \eta_p^2 = .15$; 2016-17: $\beta = -0.50, t = -3.04, p = .003; \eta_p^2 = .08$), controlling for differences on other IRI subscales. Taken together, these results hint towards relatively stable cohort differences in empathy and moral judgment, and provide no support for the influence of aging processes. However, we note that Study 2 cannot rule out the presence of a maturation effect smaller than the minimum effect size we could confidently detect.

Study 3: Time-lag analyses

In Study 1, older participants tended to make more deontological moral judgments

than younger participants, but Study 2 revealed a negligible and non-significant shift within participants over an eight-year interval. Furthermore, a Bayesian analysis lent substantial support for the absence of an aging effect.

This pattern of results opens up the possibility that cross-sectional age differences in moral judgment reflect changes in utilitarian proclivity across generations. Thus, we should observe an effect of cohort succession in a comparison between age-matched waves.

To test this prediction, in Study 3 we adopt a two-point time-lag approach. A new sample of online volunteers completed an identical study on empathy and morality between 2015 and 2017. We then apply an

exact matching (Iacus, King, Porro, & Katz, 2012) algorithm to create a 2015-17 subsample that strongly mirrors our 2007-08 sample on the primary pre-treatment covariates, thus eliminating differences in age and other potentially confounding demographic measures. In conjunction with the correlation reported in Study 1, a difference between waves would point towards a cohort effect.

Results

Imbalance correction

To evaluate whether our control (2007-08) and treatment (2015-17) samples differ in their composition, we entered every demographic predictor into a logistic regression model with wave (control vs. treatment) as the dependent measure: Participants in the 2015-17 wave were overall younger, OR = 0.98 [0.97, 0.98], more likely to be women, OR = 1.40 [1.36, 1.57], and varied also in terms of educational attainment, $\chi^2(4) = 257.70$, nationality, $\chi^2(24) = 939.83$, and religious affiliation, $\chi^2(9) = 47.25$, $ps < .001$.

Because visitors were not randomly assigned to one or the other wave, the imbalance in pre-treatment covariates may reflect bias in selection. Thus, prior to estimating our target effect, we sought to balance the distribution of covariates between waves. To this end, numerous methods for pre-processing observational data have been developed, whose general aim is to compensate for non-random assignment by either weighting or matching observations as a function of their covariate values. In our present case, given the large sample size and substantial overlap between waves on the primary demographic covariates, we were able to seek exact matches, using a *coarsened exact matching* algorithm (see the *MatchIt R package* introduced in Ho, Imai, King, & Stuart, 2011; Iacus, King, Porro, & Katz, 2012). Importantly, these techniques do not substitute for regression adjustment. Rather,

they antecede and complement statistical inference in order to reduce dependence on parametric model assumptions and thereby improve estimates of treatment effects.

Our resulting matched sample consisted of 3377 participants per wave (total $N = 6754$). The CEM algorithm achieved perfect univariate balance on educational attainment, religious affiliation, and gender (all $\chi^2s = 0$, $ps = 1$), and age was also adequately matched (age₂₀₀₇₋₀₈: $M = 28.6$, $SD = 11.8$; age₂₀₁₅₋₁₇: $M = 28.7$, $SD = 11.7$), $t(6752) = 0.36$, $p = .72$, $d = .01$. Descriptive statistics are reported in Supplementary Table 2.

Manipulation check

As a result of our matching procedure, participants in the second wave were born an average of 8.3 years later (birth-year₂₀₀₇₋₀₈ $Mdn = 1983$, $Q_1 = 1974$, $Q_3 = 1988$; birth-year₂₀₁₅₋₁₇ $Mdn = 1991$, $Q_1 = 1983$, $Q_3 = 1996$; $CLES = .73$), yielding a probability of .73 that a randomly-selected second wave participant was born after a randomly-selected first wave participant.

Cohort effects

Differences in interpersonal reactivity and religiosity between waves were all negligible, albeit significant (EC $d = -0.08$, FS $d = -0.12$, PD $d = -0.12$, religiosity $d = -0.13$, $ps < .001$, PT $d = 0.05$, $p = .03$). We did not anticipate these results: Since age differences in self-reported empathy were not related to processes of maturation, we expected a larger difference between waves. In the general discussion, we return to this issue.

As predicted, a two-sample *t*-test revealed a clear shift in moral judgment (MJ₂₀₀₇₋₀₈ = 3.91, $SD = 1.36$; MJ₂₀₁₅₋₁₇ = 4.34, $SD = 1.22$), $t(6752) = 13.6$, $d = 0.33$ [0.28, 0.37], $p < .001$, suggesting that recent cohorts in the treatment group were more likely to endorse utilitarian sacrifice than earlier cohorts in the control group. The difference between waves was significant across a range of regression models

INCREASING PERMISSIBILITY OF UTILITARIAN SACRIFICE

incorporating different sets of covariates (see Table 2).

Table 2. Multiple regression models of moral judgment after coarsened exact matching of 2007-08 (control) and 2015-17 (treatment) waves.

	Model 1 Adj. $r^2 = .130$	Model 2 Adj. $r^2 = .161$	Model 3 Adj. $r^2 = .231$
Wave (0: 2007-08, 1: 2015-17)	0.43 ** [0.37, 0.49]	0.43 ** [0.37, 0.49]	0.38 ** [0.33, 0.44]
Age	-0.03 ** [-0.03, -0.02]	-0.02 ** [-0.03, 0.02]	-0.02 ** [-0.02, -0.02]
Gender (0: male, 1: female)	-0.57 ** [-0.63, -0.51]	-0.53 ** [-0.59, -0.47]	-0.38 ** [-0.44, -0.32]
Religious affiliation			
<i>Buddhist</i>		-0.06 [-0.28, 0.15]	0.26 * [0.04, 0.47]
- <i>Catholic</i>		-0.45 ** [-0.55, -0.37]	-0.12 * [-0.22, 0.02]
- <i>Orthodox</i>		-0.65 ** [-0.90, -0.39]	-0.32 * [-0.57, -0.07]
- <i>Other</i>		-0.54 ** [-0.65, -0.43]	-0.13 * [-0.26, -0.01]
- <i>Protestant</i>		-0.45 ** [-0.55, -0.35]	-0.06 [-0.18, 0.06]
<i>Hindu</i>		-0.46 * [-0.77, -0.15]	-0.06 [-0.35, 0.25]
<i>Jewish</i>		-0.08 [-0.38, 0.22]	0.14 [-0.15, 0.43]
<i>Muslim</i>		-0.53 ** [-0.83, -0.22]	-0.06 [-0.37, 0.24]
<i>Other</i>		-0.11 [-0.24, 0.01]	0.14 * [0.01, 0.27]
Educational attainment			
<i>Bachelor's Degree</i>		-0.26 ** [-0.37, -0.15]	-0.20 ** [-0.31, -0.10]
<i>High School Graduate</i>		-0.14 * [-0.26, -0.03]	-0.12 * [-0.24, -0.01]
<i>Graduate Degree</i>		-0.22 ** [-0.34, -0.11]	-0.18 ** [-0.29, -0.07]
<i>Some College</i>		-0.20 ** [-0.30, -0.10]	-0.15 * [-0.25, -0.05]
Empathic concern			-0.36 ** [-0.40, -0.32]
Fantasy			0.10 ** [0.06, 0.14]
Perspective-taking			-0.02 [-0.06, 0.02]
Personal distress			-0.13 ** [-0.17, -0.09]
Religiosity			-0.11 ** [-0.13, -0.09]

*: $p < .05$, **: $p < .001$.

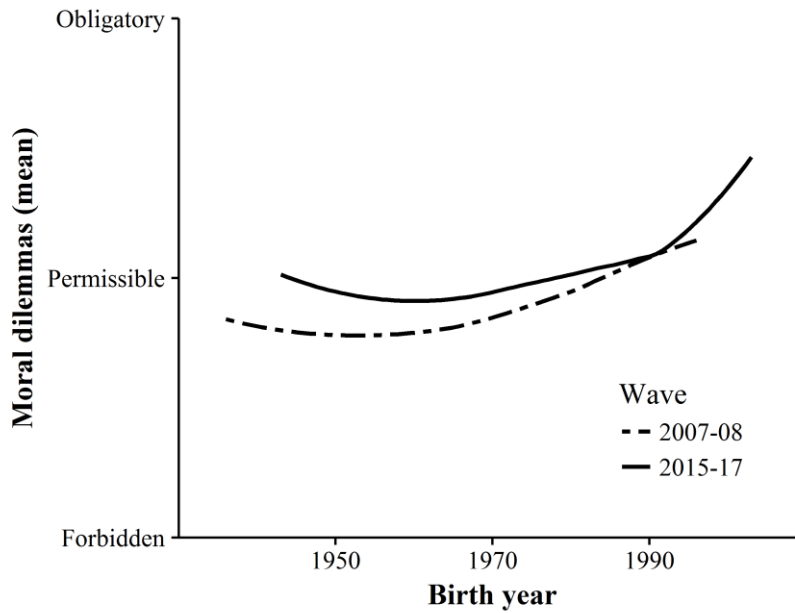


Figure 3. Moral judgment by birth year and wave, with locally-weighted (loess) smoothing.

As seen in Figure 3, the cohort increase in utilitarian judgment begins with individuals born approximately in the 1960s, and accelerates among birth cohorts after 1990. Interestingly, among earlier generations—those showing no cohort effect—an influence of period can be discerned (i.e., more utilitarian views in the second wave relative to the first), perhaps because the factors shaping the formation of recent cohorts’ utilitarian views also somewhat influence the moral psychology of mature adults.

General Discussion

In a cross-sectional design, younger participants were found to make more utilitarian judgments than older participants (Study 1). In order to understand whether this relation was primarily the consequence of aging processes that reinforce deontological intuitions or a generational shift toward utilitarian judgment, we conducted longitudinal and time-lag studies. A longitudinal approach failed to detect an effect of maturation (Study 2). Meanwhile, a comparison between two age-matched samples separated by an eight-year interval revealed increased utilitarian judgment in the recent wave (Study 3).

Together, these results are most straightforwardly interpreted by appealing to a shift in the moral psychology of recent cohorts toward greater endorsement of utilitarian sacrifice.

At the outset, we suspected that processes of secularization (Norris & Inglehart, 2011) and declining empathic concern (Konrath, O'Brien, & Hsing, 2011) might be driving change, since the opposition to welfare trade-offs has been resoundingly attributed to religious and affective prohibitions on intentional harm. Although we replicated declines in self-reported empathy and religiosity, these effects were smaller than changes in moral judgment. Below we summarize attempts to interpret this unexpected result and our recommendations for future research.

First, there is some reason to question the reliability of self-reported empathy. High trait empathy has a clear positive connotation, so the IRI subscale may be contaminated by social desirability (Watson & Morris, 1991). As a consequence, the self-report of victim-oriented feelings does not always predict empathic behavior (Devlin, Zaki, Ong, & Gruber, 2014; Marsh et al., 2014) and may

primarily serve to validate moral intuitions *post hoc* (Ditto & Liu, 2011). Instead, complementary measures of affect, using neuroimaging (Lamm, Batson & Decety, 2007) and psychophysiological (Cushman, Gray, Gaffey & Mendes, 2012) methods, may provide a clearer window into the affective processes shaping moral judgment.

Second, from a dual-process perspective, utilitarian moral judgments are often construed as the product of a cognitive suppression of affective reactions (Conway & Gawronski, 2013; Cushman, 2013; Greene, 2007; but see Bartels & Pizarro, 2011). So, growing endorsement of utilitarian sacrifice may be due, not exclusively to a reduction in the intensity of affect, but also in its relative influence upon moral judgment. Indeed, a diverse body of neuroscientific (Greene et al., 2001; Koenigs et al., 2007; Shenhav & Greene, 2014), physiological (Cushman et al., 2012; Youssef et al., 2012) and behavioral (Koenigs, Kruepke, Zeier, & Newman, 2012; Miller, Hannikainen, & Cushman, 2014; Patil, 2015) evidence now suggests that affective prohibitions of interpersonal harm are not triggered exclusively—perhaps not even primarily—by an emotional tie to the proximal victim, but also by an aversion to the harmful *act*. Nonetheless, victim concern may be an important developmental precursor to harm aversion (Cushman, 2013) and also play a decisive role in motivating helping behavior (Decety & Yoder, 2016; Habashi, Graziano, & Hoover, 2016; Hu, Strang, & Weber, 2015).

Third, scale ratings of IRI items may suffer from *recalibration* over time (Golembiewski, Billingsley, & Yeager, 1976). In contrast to our outcome measure, anchored by terms like “*obligation*” and “*prohibition*”, which may be expected to maintain relatively stable meaning over time, the self-assessment of dispositional affect may be anchored by comparison to one’s peers, confounding the test of differences between administrations.

We must also take note of a few methodological limitations. A notorious complication, known as the *identification problem*, plagues the study of long-term change. Because $age = period - cohort$, simultaneously estimating all three temporal effects on an outcome of interest yields an infinite number of solutions. Attempts to circumvent their collinearity have long been viewed with suspicion (Glenn, 1976; Luo, 2013). Instead of attempting to decompose these influences, our approach was to infer the most plausible dominant mechanism from the present pattern of results.

Independent evidence suggests that most cultural change is driven by cohort effects that withstand substantial variability over periods (Vaisey & Lizardo, 2016; also Davis, 1992). Vaisey and Lizardo (2016) demonstrate that historical change in moral attitudes and values, in particular, stems from processes of cohort replacement more so than from conversion due to contemporaneous sociocultural factors, which comports with our inference about increasing utilitarian judgment. Still, the reported findings could also emerge from the combination of a negative maturation effect and a positive period trend—an alternative we cannot decisively rule out. Discriminating between these candidate explanations requires future research following different cohorts in new periods.

In addition, most online volunteers were nationals of North American and Western European countries. We therefore remain agnostic as to whether cohort effects emerge more generally across cultures. It may be, for instance, that religious and cultural traditions vary in their baseline emphasis on deontological versus utilitarian norms and that shifts in their relative prevalence coincide somewhat with widely-studied processes of modernization and cultural change (see Inglehart, 1997). Finally, because our participants self-selected, replicating these findings using

representative sampling methods is a worthwhile task for future research.

The evolution of moral norms holds inherent interest to philosophers and social scientists (Pinker, 2011; Singer, 1981). In moral psychology, the received wisdom has been that the folk are fundamentally deontologists—opposing harmful acts carried out as a means toward the greater good (Bartels, 2008; Cushman, Young, & Hauser, 2006; Royzman & Baron, 2002). Our findings suggest, however, that this is not an immutable fact: Recent cohorts (often referred to as Millennials) are significantly more likely to support utilitarian sacrifice than their predecessors (especially Baby Boomers, born before 1970)—a divide which may contribute to the patent disagreement between younger and older adults in real-world debates about ethics and policy.

A vibrant discussion among philosophers and cognitive scientists has focused on distinguishing the virtues and pitfalls of the human moral faculty (Bloom, 2017; Greene, 2014; Singer, 2005). On a pessimistic note, our results dovetail with evidence about the socialization and development of recent cohorts (e.g., Shonkoff et al., 2012): Utilitarian judgment has been shown to correlate with Machiavellian and psychopathic traits (Bartels & Pizarro, 2011), and also with the reduced capacity to distinguish felt emotions (Patil & Silani, 2014). At the same time, leading theories credit highly-acclaimed instances of moral progress to the exercise of rational scrutiny over prevailing moral norms (Greene, 2014; Singer, 2005), and the persistence of parochialism and prejudice to the unbridled command of intuition (Bloom, 2017). From this perspective, greater disapproval of intuitive deontological principles among recent cohorts may stem from the documented rise in cognitive abilities (i.e., the Flynn effect; see Pietschnig & Voracek, 2015) and foreshadow an expanding commitment to

the welfare-maximizing resolution of contemporary moral challenges.

Future research is clearly necessary to attest to these predicted consequences of shifting moral ideologies. To this end, the present study underscores the value of epidemiological tools for moral psychology research, which enable us to detect long-term fluctuation in basic moral values.

References

- Bartels, D. M. (2008). Principled moral sentiment and the flexibility of moral judgment and decision making. *Cognition*, *108*(2), 381-417.
- Bartels, D. M., & Pizarro, D. A. (2011). The mismeasure of morals: Antisocial personality traits predict utilitarian responses to moral dilemmas. *Cognition*, *121*(1), 154-161.
- Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E.-J., Berk, R., Bollen, K. A., Brembs, B., Brown, L., Camerer, C., Cesarini, D., Chambers, C. D., Clyde, M., Cook, T. D., De Boeck, P., Dienes, Z., Dreber, A., Easwaran, K., Efferson, C., Fehr, E., Fidler, F., Field, A. P., Forster, M., George, E. I., Gonzalez, R., Goodman, S., Green, E., Green, D. P., Greenwald, A., Hadfield, J. D., Hedges, L. V., Held, L., Ho, T.-H., Hoijtink, H., Jones, J. H., Hruschka, D. J., Imai, K., Imbens, G., Ioannidis, J. P. A., Jeon, M., Kirchler, M., Laibson, D., List, J., Little, R., Lupia, A., Machery, E., Maxwell, S. E., McCarthy, M., Moore, D., Morgan, S. L., Munafó, M., Nakagawa, S., Nyhan, B., Parker, T. H., Pericchi, L., Perugini, M., Roulder, J., Rousseau, J., Savalei, V., Schönbrodt, F. D., Sellke, T., Sinclair, B., Tingley, D., Van Zandt, T., Vazire, S., Watts, D. J., Winship, C., Wolpert, R. L., Xie, Y., Young, C., Zinman, J., & Johnson, V. E. (2017). Redefine statistical significance. *Nature Human Behavior*. doi:10.1038/s41562-017-0189-z.
- Bloom, P. (2017). *Against empathy: The case for rational compassion*. Random House.
- Conway, P., & Gawronski, B. (2013). Deontological and utilitarian inclinations in moral decision making: a process

- dissociation approach. *Journal of Personality and Social Psychology*, 104(2), 216.
- Cushman, F. (2013). Action, outcome, and value a dual-system framework for morality. *Personality and Social Psychology Review*, 17(3), 273-292.
- Cushman, F., Gray, K., Gaffey, A., & Mendes, W. B. (2012). Simulating murder: the aversion to harmful action. *Emotion*, 12(1), 2.
- Cushman, F., Young, L., & Hauser, M. (2006). The role of conscious reasoning and intuition in moral judgment: Testing three principles of harm. *Psychological Science*, 17(12), 1082-1089.
- Davis, J. A. (1992). Changeable weather in a cooling climate atop the liberal plateau: Conversion and replacement in forty-two General Social Survey items, 1972–1989. *Public Opinion Quarterly*, 56(3), 261-306.
- Decety, J., & Yoder, K. J. (2016). Empathy and motivation for justice: Cognitive empathy and concern, but not emotional empathy, predict sensitivity to injustice for others. *Social neuroscience*, 11(1), 1-14.
- Devlin, H. C., Zaki, J., Ong, D. C., & Gruber, J. (2014). Not as good as you think? Trait positive emotion is associated with increased self-reported empathy but decreased empathic performance. *PloS one*, 9(10), e110470.
- Ditto, P. H., & Liu, B. (2011). Deontological dissonance and the consequentialist crutch. *The social psychology of morality: Exploring the causes of good and evil*, 51-70.
- Foot, P. (1967). The problem of abortion and the doctrine of double effect.
- Gleichgerrcht, E., & Young, L. (2013). Low levels of empathic concern predict utilitarian moral judgment. *PloS one*, 8(4), e60418.
- Glenn, N. D. (1976). Cohort analysts' futile quest: Statistical attempts to separate age, period and cohort effects. *American Sociological Review*, 41(5), 900-904.
- Golembiewski, R. T., Billingsley, K., & Yeager, S. (1976). Measuring change and persistence in human affairs: Types of change generated by OD designs. *The Journal of Applied Behavioral Science*, 12(2), 133-157.
- Greene, J. D. (2007). The secret joke of Kant's soul. *Moral psychology: Historical and contemporary readings*, 359-372.
- Greene, J. D. (2014). *Moral tribes: Emotion, reason and the gap between us and them*. Atlantic Books Ltd.
- Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, 293(5537), 2105-2108.
- Habashi, M. M., Graziano, W. G., & Hoover, A. E. (2016). Searching for the prosocial personality: A Big Five approach to linking personality and prosocial behavior. *Personality and Social Psychology Bulletin*, 42(9), 1177-1192.
- Hannikainen, I., Miller, R., & Cushman, F. (2017). Act versus Impact: Conservatives and liberals exhibit different structural emphases in moral judgment. *Ratio*, advance online publication. doi:10.1111/rati.12162
- Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2007). Match It: Nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software*, 42(8), 1-28.
- Hu, Y., Strang, S., & Weber, B. (2015). Helping or punishing strangers: neural correlates of altruistic decisions as third-party and of its relation to empathic concern. *Frontiers in behavioral neuroscience*, 9.
- Iacus, S. M., King, G., Porro, G., & Katz, J. N. (2012). Causal inference without balance checking: Coarsened exact matching. *Political Analysis*, 1-24.
- Inglehart, R. (1997). *Modernization and postmodernization: Cultural, economic, and political change in 43 societies*. Princeton University Press.
- Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., & Damasio, A. (2007). Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature* 446:908–911.
- Koenigs, M., Kruepke, M., Zeier, J., & Newman, J. P. (2012). Utilitarian moral judgment in psychopathy. *Social Cognitive and Affective Neuroscience*, 7(6), 708-714.
- Konrath, S. H., O'Brien, E. H., & Hsing, C. (2011). Changes in dispositional empathy in American college students over time: A meta-analysis. *Personality and Social Psychology Review*, 15(2), 180-198.

- Lamm, C., Batson, C. D., & Decety, J. (2007). The neural substrate of human empathy: effects of perspective-taking and cognitive appraisal. *Journal of Cognitive Neuroscience*, *19*(1), 42-58.
- Luo, L. (2013). Assessing validity and application scope of the intrinsic estimator approach to the age-period-cohort problem. *Demography*, *50*(6), 1945-1967.
- Marsh, A. A., Stoycos, S. A., Brethel-Haurwitz, K. M., Robinson, P., VanMeter, J. W., & Cardinale, E. M. (2014). Neural and cognitive characteristics of extraordinary altruists. *Proceedings of the National Academy of Sciences*, *111*(42), 15036-15041.
- Miller, R., Hannikainen, I., & Cushman, F. (2014). Bad actions or bad outcomes? Differentiating affective contributions to the moral condemnation of harm. *Emotion*, *14*(3), 573.
- Norris, P., & Inglehart, R. (2011). *Sacred and secular: Religion and politics worldwide*. Cambridge University Press.
- O'Brien, E., Konrath, S. H., Gröhn, D., & Hagen, A. L. (2013). Empathic concern and perspective taking: Linear and quadratic effects of age across the adult life span. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *68*(2), 168-175.
- Patil, I. (2015). Trait psychopathy and utilitarian moral judgement: The mediating role of action aversion. *Journal of Cognitive Psychology*, *27*(3), 349-366.
- Patil, I., & Silani, G. (2014). Reduced empathic concern leads to utilitarian moral judgments in trait alexithymia. *Frontiers in Psychology*, *5*, 501.
- Piazza, J., & Landy, J. F. (2013). "Lean not on your own understanding": Belief that morality is founded on divine authority and non-utilitarian moral judgments. *Judgment and Decision Making*, *8*(6), 639.
- Pietschnig, J., & Voracek, M. (2015). One century of global IQ gains: A formal meta-analysis of the Flynn effect (1909–2013). *Perspectives on Psychological Science*, *10*(3), 282-306.
- Pinker, S. (2011). *The better angels of our nature: The decline of violence in history and its causes*. Penguin UK.
- Royzman, E. B., & Baron, J. (2002). The preference for indirect harm. *Social Justice Research*, *15*(2), 165-184.
- Schwadel, P. (2010). Period and cohort effects on religious nonaffiliation and religious disaffiliation: A research note. *Journal for the Scientific Study of Religion*, *49*(2), 311-319.
- Shenhav, A., & Greene, J. D. (2014). Integrative moral judgment: dissociating the roles of the amygdala and ventromedial prefrontal cortex. *Journal of Neuroscience*, *34*(13), 4741-4749.
- Shenhav, A., Rand, D. G., & Greene, J. D. (2012). Divine intuition: cognitive style influences belief in God. *Journal of Experimental Psychology: General*, *141*(3), 423.
- Shonkoff, J. P., Garner, A. S., Siegel, B. S., Dobbins, M. I., Earls, M. F., McGuinn, L., ... & Committee on Early Childhood, Adoption, and Dependent Care. (2012). The lifelong effects of early childhood adversity and toxic stress. *Pediatrics*, *129*(1), e232-e246.
- Singer, P. (1981). *The expanding circle*. Oxford: Clarendon Press.
- Singer, P. (2005). Ethics and intuitions. *The Journal of Ethics*, *9*(3-4), 331-352.
- Vaisey, S., & Lizardo, O. (2016). Cultural fragmentation or acquired dispositions? A new approach to accounting for patterns of cultural change. *Socius*, *2*, 2378023116669726.
- Watson, P. J., & Morris, R. J. (1991). Narcissism, empathy and social desirability. *Personality and Individual Differences*, *12*(6), 575-579.
- Youssef, F. F., Dookeeram, K., Basdeo, V., Francis, E., Doman, M., Mamed, D., ... & Legall, G. (2012). Stress alters personal moral decision making. *Psychoneuroendocrinology*, *37*(4), 491-498.