Who Shalt Not Kill? Individual Differences in Working Memory Capacity, Executive Control, and Moral Judgment

By: Adam B. Moore, Brian A. Clark, and Michael J. Kane

Moore, A.B., Clark, B.A., & Kane, M.J. (2008). Who shalt not kill? Individual differences in working memory capacity, executive control, and moral judgment. *Psychological Science*, *19*, 549-557.

Abstract:

Recent findings suggest that exerting executive control influences responses to moral dilemmas. In our study, subjects judged how morally appropriate it would be for them to kill one person to save others. They made these judgments in 24 dilemmas that systematically varied physical directness of killing, personal risk to the subject, inevitability of the death, and intentionality of the action. All four of these variables demonstrated main effects. Executive control was indexed by scores on working-memory-capacity (WMC) tasks. People with higher WMC found certain types of killing more appropriate than did those with lower WMC and were more consistent in their judgments. We also report interactions between manipulated variables that implicate complex emotion-cognition integration processes not captured by current dual-process views of moral judgment.

Article:

> When Pilate saw that he could prevail nothing, but that rather a tumult was made, he took water, and washed his hands before the multitude, saying, I am innocent of the blood of this just person.... (Matthew 27:24, King James Version)

The synoptic gospels are notoriously ambiguous in assigning earthly culpability for Jesus' death. Rather than actively sentencing Jesus to crucifixion, Pilate passively allows it by letting the multitude decide whether to release Jesus or Barabbas. When the crowd chooses Barabbas, Pilate claims his hands are clean regarding Jesus' fate. We suggest that the gospel authors captured something important about humankind's moral sense, particularly regarding the morality of being the direct agent of killing.

Philosophers have long considered such issues in developing prescriptive ethics, but scientists are only now studying moral judgments in order to understand the generative processes that influence them. A lively theoretical debate concerns the role of emotion versus that of normative-rule-based cognition in driving moral judgments (e.g., Blair, 1995; Haidt, 2001; Lakoff, 2002; Mikhail, 2007; Nichols, 2002; Pizarro & Bloom, 2003; Prinz, 2006). Much of the excitement surrounding this debate derives from the neuroimaging work of Greene and his colleagues (Greene, Nystrom, Engell, Darley, & Cohen, 2004; Greene, Sommerville, Nystrom, Darley, & Cohen, 2001; see also Borg, Hynes, Van Horn, Grafton, & Sinnott-Armstrong, 2006). Using functional magnetic resonance imaging (fMRI), Greene et al. investigated the neural substrates of judgments about hypothetical personal versus impersonal dilemmas. In personal dilemmas, subjects contemplated (a) causing serious bodily harm or death to (b) another person or persons, (c) in a way that did not simply deflect preexisting harm onto another person; personal dilemmas thus involved "up close and personal" harmful acts the subjects directly initiated. Impersonal dilemmas did not meet these three criteria.

These authors illustrated personal dilemmas with the footbridge problem, in which a runaway trolley hurtles toward five unaware workmen; the only way to save them is to push a heavy man (standing nearby on a footbridge) onto the track, where he will die in stopping the trolley. In an impersonal version of this dilemma, the trolley problem, one may save the workmen by throwing a switch that simply diverts the trolley onto another track, killing one worker. Most people find the action in the impersonal dilemma to be morally more acceptable than the action in the personal dilemma, despite their identical consequences (e.g., Hauser, Cushman,

Young, Jin, & Mikhail, 2007). Why? Emotion prevails in personal dilemmas. Greene et al. (2001, 2004) demonstrated that personal dilemmas engage brain regions involved in emotion (e.g., amygdala, posterior cingulate/ precuneus), whereas impersonal dilemmas activate areas involved in deliberative reasoning and working memory (e.g., middle frontal gyrus, bilateral parietal lobe). Moreover, when subjects judge resolutions to personal dilemmas to be morally appropriate, they respond slowly and engage networks associated with executive control (e.g., anterior dorsolateral prefrontal cortex, anterior cingulate). Greene et al. thus proposed a dual-process model: Personal moral dilemmas evoke an automatic, "hot" emotional response that biases against responding in a way that causes harm, but this response may be overridden by "cold" cognitive control.

In our study, we tested this dual-process theory by asking whether individual differences in control, indexed by scores on working-memory-capacity (WMC) tasks, would predict cold, consequentialist responses to personal dilemmas. Research with nonemotive tasks indicates that (a) WMC variation predicts executive ability to override prepotent responses (e.g., Kane, Bleckley, Conway, & Engle, 2001; Kane & Engle, 2003), (b) WMC variation predicts abstract-reasoning performance (e.g., Kane, Hambrick, & Conway, 2005; Oberauer, Schulze, Wilhelm, & SüB, 2005), and (c) experimentally increasing working memory load impairs reasoning about future consequences (e.g., Hinson, Jameson, & Whitney, 2003). Therefore, we hypothesized that by controlling emotion and engaging deliberative processing, people with greater WMC would more rationally evaluate consequences within personal moral dilemmas. This is a strong prediction because executive theories of WMC (e.g., Engle & Kane, 2004; Hasher, Lustig, & Zacks, 2007) are supported almost exclusively by studies of cold cognition (but see Braver, Gray, & Burgess, 2007), and individual differences influencing such moral judgments are rare (Hauser et al., 2007; O'Neill & Petrinovich, 1998).

Our second goal was to test whether the personal-impersonal distinction would survive a retooling of the dilemmas presented to subjects.[1] To do this, we redesigned the materials used by Greene et al. (2001, 2004) to address several significant problems with them: (a) More personal than impersonal dilemmas involved death or injury, and more impersonal than personal dilemmas involved lying or stealing; (b) unlike personal dilemmas, many impersonal dilemmas presented abstract, probabilistic-reasoning problems; (c) several scenarios posed nondilemmas (e.g., should a child murder his grandmother for not buying him a gift?); (d) the descriptions of personal dilemmas were longer than the descriptions of impersonal dilemmas (Ms = 124 and 102 words, respectively); (e) whether or not subjects' lives were endangered was varied unsystematically across the dilemmas; and (f) subjects saw multiple versions of some scenarios, so carryover effects were possible (a potential problem for other, similar studies as well: Hauser et al., 2007; Mikhail, 2007; Petrinovich, O'Neill, & Jorgensen, 1993; Valdesolo & DeSteno, 2006).

Some of these shortcomings reflect vagueness in how Greene et al. (2001, 2004) defined impersonal dilemmas. Personal dilemmas met three criteria (bodily harm that was caused to persons and did not simply deflect existing harm), and any dilemmas meeting none, one, or two of these criteria were considered impersonal. Consequently, very different dilemmas were contrasted, so the materials failed to systematically capture the vexing psychological distinction between the footbridge and trolley problems. For example, dilemmas involving whether to make charitable donations, steal or damage property, or enact environmental-hazard policies were included in the impersonal dilemmas along with those involving deflecting impending threats of death (e.g., a runaway trolley, poisonous fumes) from an innocent group toward an innocent individual. To better equate the severity of outcomes of personal and impersonal dilemmas, we redefined the personal-impersonal distinction to reflect more versus less direct killing (see also Royzman & Baron, 2002; Spranca, Minsk, & Baron, 1991). All critical dilemmas involved killing one individual to save more people, but the killing contemplated in personal dilemmas was more physically direct than the killing contemplated in impersonal dilemmas; that is, the killing involved in personal dilemmas was less mediated through mechanical-technological means or through other people's actions.

Our third goal was to test whether other philosophically relevant variables influence moral judgment (perhaps moderating any effects of the personal-impersonal variable). First, because everyday moral dilemmas often involve costs and benefits to the self, and trolley-dilemma judgments change with subjects' relationships to the

hypothetical victims (Petrinovich et al., 1993), we systematically varied self-risk. We predicted that killing to save oneself and other people would be more acceptable than killing to save only other people; we view this self-other distinction as evolutionarily ingrained and thus automatically and intuitively processed (cf. Petrinovich et al.). Second, we manipulated whether the intended object of harm would die regardless of subjects' actions, and we predicted that subjects would find killing more acceptable if it merely hastened an inevitable death than if the death were avoidable; we view this inevitable-avoidable distinction as relatively subtle and thus more rationally and deliberatively processed than the self-other distinction. Third, as have other researchers (Borg et al., 2006; Cushman, Young, & Hauser, 2006; Hauser et al., 2007; Mikhail, 2007), we examined Thomas Aquinas' (trans. 1988, pp. 226– 227) doctrine of double effect (DDE), according to which harm is more permissible if it is a foreseen but unintended consequence than if it is an intended means to an end. Normative sensitivity to the DDE may partially explain the effects of the personal-impersonal distinction: In footbridge/trolley-type dilemmas, one either intentionally uses the bystander's body (e.g., by throwing him off the footbridge) or unintentionally sacrifices another workman (e.g., by throwing the switch). People may therefore blanch at resolutions to personal dilemmas because they have internalized the DDE. Thus, half of our impersonal dilemmas required killing as a means, and half required killing as an unintended but foreseeable outcome.

METHOD

Subjects

Anticipating future neuroimaging research on this topic, we used strict inclusion criteria. Specifically, subjects had to report no history of psychiatric diagnoses and no current illness or (psychiatric or flu) medication for their data to be included in the analyses. Of 145 subjects who completed WMC screening and the moral-judgment task in exchange for partial credit toward an undergraduate course requirement, 32 had their data excluded: 16 who had a previous history of head trauma or were currently taking psychiatric or flu medication, 12 who ignored instructions, 3 whose WMC data were invalid, and 1 who was agitated. We report data from 113 native-English-speaking subjects (68 females, 45 males).

WMC Screening

Subjects individually completed three automated complex-span measures: operation span (OSPAN), reading span (RSPAN), and symmetry span (SSPAN). In each case, memory items were interleaved with items from a processing task. The response deadline for each processing task was tailored to that particular task and subject, based on latencies (M + 2.5 SD) obtained during pretesting with 15 processing-only items (Unsworth, Heitz, Schrock, & Engle, 2005). The processing tasks were verifying equations involving two operations (OSPAN), verifying the meaningfulness of sentences (RSPAN), and judging matrix patterns for symmetry (SSPAN). Each processing stimulus was presented until the subject responded or the deadline was reached, and was followed 200 ms later by a memory item (presented for 250 ms in OSPAN and RSPAN and for 650 ms in SSPAN); each memory item was followed by the next processing stimulus or a memory test.

In the OSPAN and RSPAN tasks, subjects were required to remember lists of 3, 4, 5, 6, or 7 letters (taken from a pool of 12) in serial order; in the SSPAN task, subjects were required to remember the sequential locations of 2, 3, 4, or 5 red squares presented within an empty 4 x 4 matrix. For OSPAN and RSPAN, each memory test presented the pool of 12 possible letters (in the same locations each time), and subjects used a computer mouse to click on the memory items in serial order. Each SSPAN test item presented an empty 4 x 4 matrix, and subjects clicked the previously occupied (red) squares in serial order. For each task, three trials at each list length were presented in random order.

Span scores (total number of items recalled in the correct serial position; Conway et al., 2005) correlated with one another, rs = .70 (OSPAN-RSPAN), .54 (OSPAN-SSPAN), and .64 (RSPAN-SSPAN). We created a WMC score for each subject by averaging z scores for the three tasks; WMC scores were normally distributed (skew = -0.497, kurtosis = -0.452).

Moral-Judgment Task

Materials

We created dilemmas based on the contextual and consequential equivalence exemplified by the footbridge (personal) and trolley (impersonal) scenarios of Greene et al. (2001, 2004), modifying their materials whenever possible into critical or filler dilemmas, but also inventing our own (for examples, see Table 1; a complete set of the materials is available on the Web at http://www.uncg. edu/—mjkane/memlab.html). The subject was the protagonist in all dilemmas, each of which consisted of an introductory paragraph and a four- or five-sentence resolution. All subjects saw the same 14 fillers, which were similar to the critical dilemmas except that 12 involved no deaths, 1 called for killing many people to save one individual, and 1 involved killing for personal gain. The 14 fillers made the content of our stimulus list resemble that of Greene et al. (2001, 2004), while controlling for confounds (as discussed in the following paragraphs).

Each of our 24 critical scenarios had two resolutions: personal and impersonal. Subjects saw only one version of each dilemma (counterbalanced across subjects). All critical dilemmas involved killing at least one person in order to save more, pitting the desire to maximize good consequences against that to avoid causing harm. Wherever possible, we used identical phrasing for personal and impersonal resolutions, equated them for the number of deaths, and matched the length of the versions to within two words.

Subjects' lives were at risk in 12 of the 24 critical dilemmas, so that killing saved oneself and others (*self* dilemmas); subjects were not in jeopardy in the other 12 (*other* dilemmas). Additionally, in 12 of the 24 critical dilemmas, the sacrificed life would be lost regardless of the subject's actions (*inevitable* dilemmas); in the other 12, the life would be lost only if the subject acted (*avoidable* dilemmas). Our materials reflected a 2 x 2 x 2 nested design, with three personal and three impersonal resolutions each nested within each of the 6 self-inevitable, 6 self-avoidable, 6 other-inevitable, and 6 other-avoidable dilemmas. Mean sentence and word counts were matched (e.g., M_{words} = 100.9, 95.3, 93.3, and 101.2 for self-inevitable, self-avoidable, other-inevitable, and other-avoidable dilemmas, respectively). We created two lists, such that each dilemma was presented with one resolution (personal or impersonal) in one list and with the complementary resolution in the other. To test the influence of the DDE on judgments, we constructed the impersonal resolutions so that 6 required killing one person as a means to save more people (*instrumental* dilemmas) and 6 involved actions that led to the death of one person as a foreseen but unintended consequence (*incidental* dilemmas). All personal dilemmas were instrumental.

Procedure

We tested all subjects individually. Dilemmas were presented on a computer monitor, in black type against a gray background. The experimenter read aloud on-screen instructions encouraging subjects to disregard legality and consider only moral appropriateness. Subjects completed two practice dilemmas before beginning the experimental trials. We randomized the order in which dilemmas were presented for each subject.

Each introductory paragraph appeared alone until subjects pressed a key to begin viewing the resolution sentences; with each key press, an additional sentence was revealed, until all were in view (all text remained visible until the final response was made). The final sentence asked whether the resolution was morally appropriate; subjects responded by pressing one of two keys, using one hand for each key. The screen went blank for 1 s between dilemmas. After the experiment, subjects completed a demographic-health questionnaire that provided the information we needed to apply the previously mentioned inclusion criteria.

RESULTS

All null-hypothesis significance tests were nondirectional, with an alpha level of .05. We report η_p^2 as our measure of effect size; p_{rep} indicates the probability of replicating the direction of the effect given similar procedures and subjects (Killeen, 2005).

TABLE 1

Text of Sample Dilemmas

Condition	Scenario	Resolution
Self-benefit, avoidable death	You and five other people are trapped in a burning building. There is only one emergency exit through which all of you could escape to safety, but it is blocked by burning debris. You notice another person in the hallway leading to the exit who has been injured but is about to crawl to safety through a small hole at the bottom of the exit door. You and the five people behind you do not have time to climb through the small hole.	(a) *Personal resolution*: You realize that you could grab the injured man and use his body as a battering-ram to break through the burning blockage in the hallway that is preventing your escape. Doing this is certain to kill him. However, if you do not do this, you and the five people behind you will die. Is it appropriate for you to kill this man in order to save yourself and the other five people? (b) *Impersonal resolution*: The hallway's emergency system puts out fire by eliminating oxygen from the hall and you can activate the system by pressing a nearby button. The fire will go out, but the injured person will suffocate and die. However, if you do not do this, you and the five people behind you will die. Is it appropriate for you to activate the system in order to save yourself and the other five people?
Other-benefit, inevitable death	You are a worker in a mine. The only way to exit the mine is to ride up in rock-buckets that can hold up to three people at a time. It is the end of the workday and miners from lower levels are riding up past you. As you are waiting for a bucket to ride in, you notice that the cable supporting the rock buckets is about to snap.	(a) *Personal resolution:* If the cable snaps, all of the miners in the buckets will fall to their deaths. The only way to prevent this is to use your axe to hit the last bucket causing it to flip over and dump its contents, lightening the load enough to save the miners above. There is one miner in this bucket who will be killed as a result. Is it appropriate to flip over the last miner's bucket, killing him, to save the other miners? (b) *Impersonal resolution:* If the cable snaps, all of the miners in the buckets will fall to their deaths. The only way to prevent this is to hit the emergency bucket release switch which will automatically detach the last bucket from the cable, lightening the load just enough to save the miners above. There is one miner in this bucket who will be killed as a result. Is it appropriate to detach the last miner's bucket, killing him, to save the other miners?

Responses to the Dilemmas

The mean proportions of affirmative responses (i.e., that killing was morally appropriate) showed the predicted pattern of results for the personal-impersonal, self-other, and inevitable-avoidable variables (see Fig. 1). A repeated measures analysis of variance (ANOVA) confirmed that impersonal killing was judged to be more appropriate than personal killing, $F(1,112) = 48.05$, $\eta_p^2 = .30$, $p_{rep} > .99$; killing to save oneself and others was judged to be more appropriate than killing to save only others, $F(1, 112) = 16.88$, $\eta_p^2 = .13$, $p_{rep} > .99$; and killing someone whose death was inevitable was judged to be more appropriate than killing someone whose death was avoidable, $F(1, 112) = 31.35$, $\eta_p^2 = .22$, $p_{rep} > .99$.[2] The only interaction among these variables that was highly replicable was their three-way interaction, $F(1, 112) = 8.29$, $\eta_p^2 = .07$, $p_{rep} = .98$.

We first explored this interaction with a 2 (personal, impersonal) x 2 (inevitable, avoidable) ANOVA on the self dilemmas. The interaction, $F(1, 112) = 4.13$, $\eta_p^2 = .04$, $p_{rep} = .96$, demonstrated that the personal-impersonal distinction had a stronger effect for inevitable deaths than for avoidable deaths: When subjects' own lives were at risk in the scenario, impersonal killing was judged to be more appropriate if victims' deaths were inevitable than if those deaths were avoidable, $t(112) = 5.04$, η_p^2 5 .185, $p_{rep} > .99$; personal killing did not elicit this effect, $t(112) = 1.23$, $p_{rep} = .81$. The parallel ANOVA on other (i.e., non-self) dilemmas yielded the opposite pattern, $F(1, 112) = 5.24$, $\eta_p^2 = .05$, p_{rep} 5 .94. When subjects' own lives were not at risk, impersonal killing was judged to be equally appropriate regardless of whether victims' deaths were inevitable or avoidable, $t(112) = 1.61$, $p_{rep} = .87$, but personal killing was judged to be more acceptable when the deaths were inevitable, $t(112)$ 5 4.72, Z_p^2 5 .17, $p_{rep} > .99$. Note also that judgments in the two personal-self conditions and in the personal-other-inevitable condition did not differ (all ts < 1.23, all p_{rep}s < .81), indicating that although either self-threat

or inevitability of death, alone, attenuated aversion to personal killing, they had no combined effect. In contrast, when impersonal killing was required, the self-other and inevitable-avoidable variables interacted to yield particularly high "appropriate" judgments in the self-inevitable condition; that is, the proportion of "appropriate" judgments was higher in the impersonal-self-inevitable condition than in the impersonal-self-avoidable and the impersonal-other-inevitable conditions, ts(112) > 3.65, p_{rep}s > .99, and the proportions in the latter two conditions were equivalent, t(112) = 0.419, prep = .62.

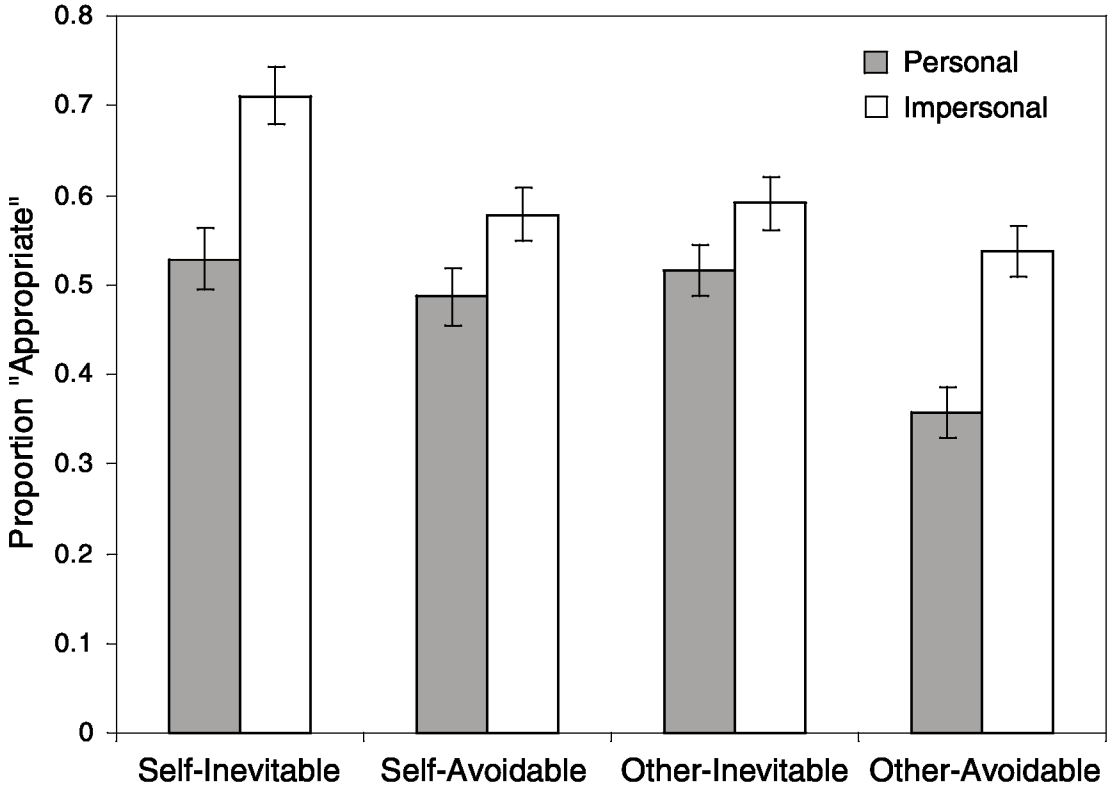

Fig. 1. Mean proportion of "appropriate" responses to the resolutions of moral dilemmas as a function of dilemma type (personal vs. impersonal harm, self vs. other benefit, and inevitable vs. avoidable death). Error bars depict standard errors.

To address the influence of the DDE, we tested impersonal dilemmas for effects of instrumental versus incidental killing. Impersonal killing was judged to be more appropriate when it involved the unintentional but foreseen death of another person than when it involved killing as an intentional means to an end (Ms = .635 and .585, respectively), t(112) = 2.33, η_p^2 = .05, p_{rep} = .95. However, impersonal instrumental killing was judged to be more appropriate than was personal killing overall (Ms = .585 and .472, respectively), t(112) = 4.79, p_{rep} > .99; thus, the DDE cannot account for the effect of the personal-impersonal distinction on moral judgments.

Role of WMC

We assessed the role of WMC in moral judgment by reconducting our analyses with WMC as a covariate (treating WMC as a continuous variable; Conway et al., 2005; Oberauer, 2005). The only effect involving WMC was its interaction with the personal-impersonal and inevitable-avoidable variables, F(1, 111) = 3.43, η_p^2 = .03, p_{rep} = .90. Figure 2 presents the mean proportions of "appropriate" judgments (collapsed over the self-other variable) for subjects with low, medium, and high WMC (defined by tertiary split); the WMC groups' judgments appeared to differ only for personally hastening inevitable deaths.

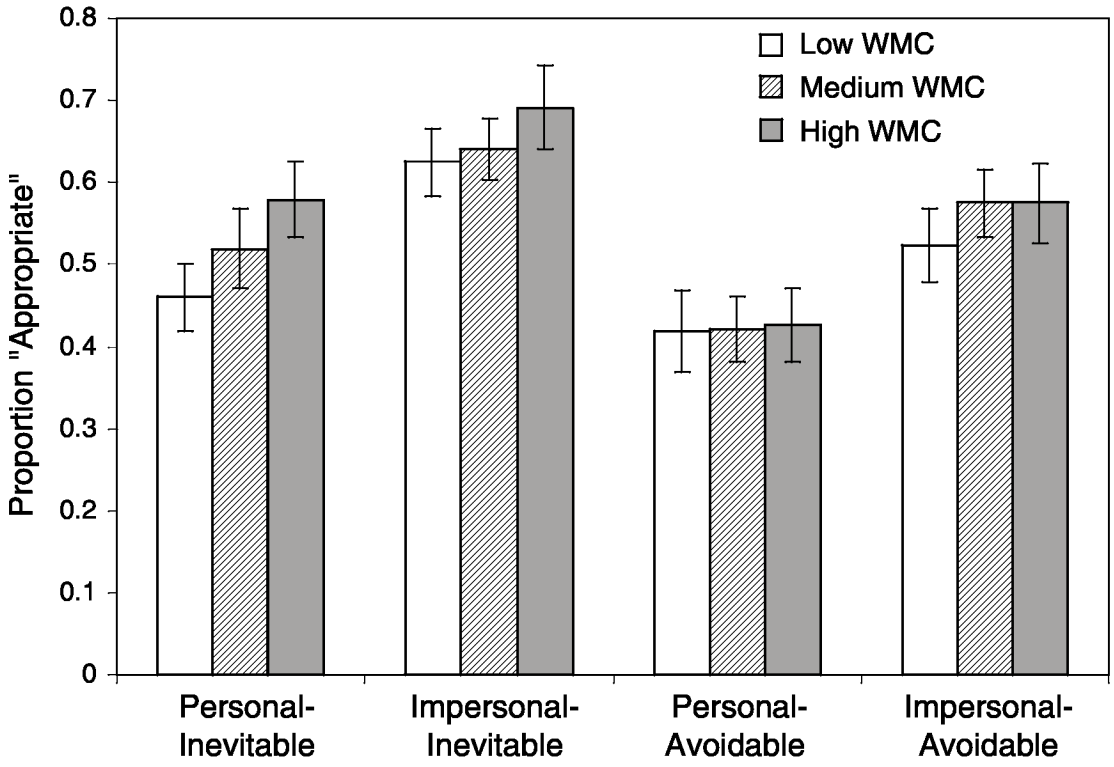

Fig. 2. Mean proportion of "appropriate" responses to the resolutions of moral dilemmas as a function of dilemma type (personal vs. impersonal harm and inevitable vs. avoidable death) and working memory capacity (WMC). Error bars depict standard errors.

We pursued this interaction with an analysis of covariance (ANCOVA) for personal dilemmas, which yielded a WMC x Inevitable-Avoidable interaction, $F(1, 111) = 3.28$, $\eta_p^2 = .03$, $p_{rep} = .90$. When considering killing an individual who would die anyway, subjects with higher WMC scores found personal killing to be more appropriate than did subjects with lower WMC scores; subjects in the three WMC groups found personal killing equally inappropriate for avoidable deaths. The parallel ANCOVA on impersonal dilemmas found no effects of WMC, $Fs < 1$. Thus, when judging the appropriateness of personal killing, only subjects with higher WMC altered their responding in light of inevitability of the victim's death. (An ANCOVA testing the DDE by analyzing the effects of the impersonal-instrumental and impersonal-incidental variables found no effects of WMC, $Fs<1$.)

A final WMC analysis tested whether better cognitive control predicted greater consistency of judgment, especially for personal dilemmas, which presumably challenged cognitive control more strongly than did impersonal dilemmas. For each subject, we scored each "appropriate" response as 1 and each "inappropriate" response as 0, and then calculated the standard deviation across judgments. Because WMC interacted with only the personal-impersonal and inevitable-avoidable variables, we collapsed standard deviations over the self-other variable. Our ANCOVA results indicated that judgments of subjects with higher WMC scores were less variable than were those of subjects with lower WMC scores, $F(1,111) = 4.29$, $\eta_p^2 = .04$, $p_{rep} = .93$, and that WMC affected variability differently in personal and impersonal dilemmas, $F(1, 111) = 3.48$, $\eta_p^2 = .03$, $p_{rep} = .90$ (for other WMC effects, $Fs < 1$). Follow-up analyses verified that WMC affected consistency across personal dilemmas, $F(1,111) = 7.55$, $\eta_p^2 = .064$, $p_{rep} = .97$, but not across impersonal dilemmas ($F < 1$; see Fig. 3).

Response Time

Results from a 2 (personal, impersonal) x 2 (self, other) x 2 (inevitable, avoidable) repeated measures ANOVA[3] on mean response times (RTs) indicated that personal dilemmas were judged faster than impersonal dilemmas (Ms = 4,783 and 5,601 ms, respectively), $F(1,110) = 18.62$, $\eta_p^2 = .15$, $p_{rep} > .99$, but the self-other and inevitable-avoidable variables had no main effects (Fs < 1). The interaction of the personal-impersonal and inevitable-avoidable variables was significant, $F(1, 110) = 4.18$, $\eta_p^2 = .04$, $p_{rep} = .92$. For personal killing, inevitable deaths yielded slower judgments than did avoidable ones (Ms = 4,999 and 4,566 ms, respectively), $t(110) = 2.29$, $\eta_p^2 = .05$, $p_{rep} = .94$; this result suggests that some deliberation was required to reconcile a bias against personal killing with that toward hastening inevitable death. For impersonal killing, inevitable and avoidable scenarios yielded equivalent RTs (Ms = 5,415 and 5,788 ms, respectively), $t(110) = -1.22$, $p_{rep} = .81$. RTs did not differ between impersonal-instrumental and impersonal-incidental killing (Ms = 5,844 and 6,221 ms, respectively; 5 subjects were excluded because of empty cells), $F(1, 107) = 1.44$, $p_{rep} = .80$; thus, RTs did not show an influence of the DDE.

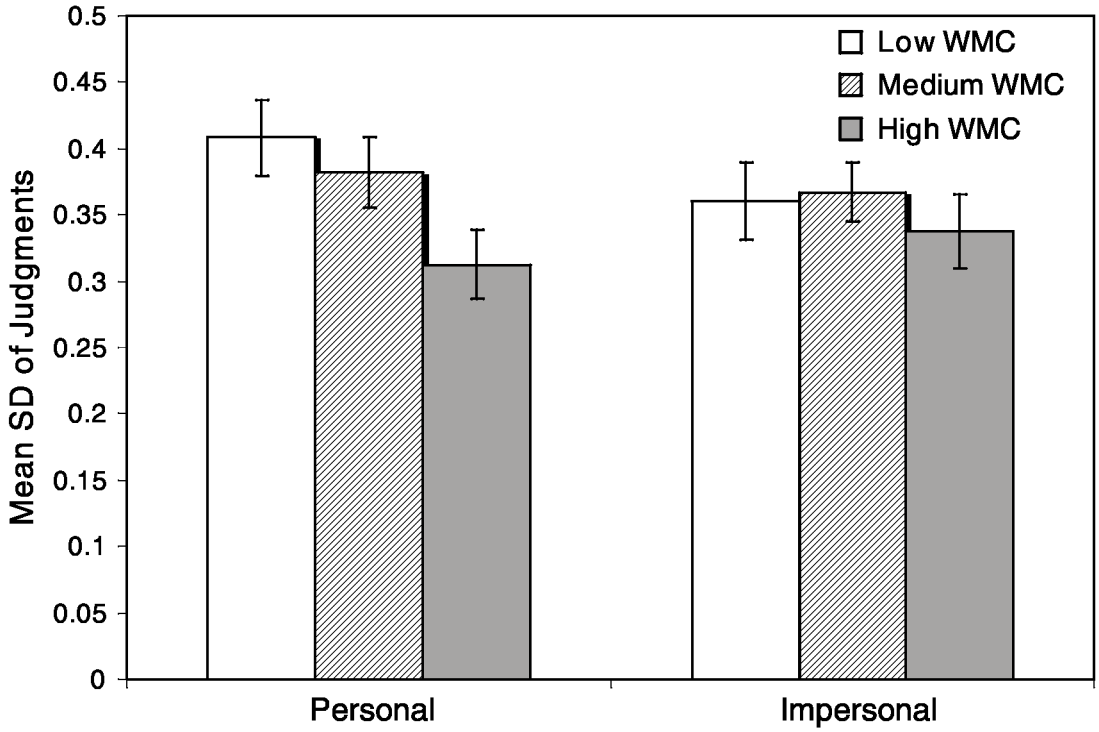

Fig. 3. Mean standard deviation across moral judgments as a function of dilemma type (personal vs. impersonal) and working memory capacity (WMC). Error bars depict standard errors.

We also examined whether our data replicated the Stroop-like finding (Greene et al., 2001) of slowed "appropriate" responses to personal dilemmas relative to impersonal dilemmas (and relative to "inappropriate" responses to personal dilemmas). To do so, we treated response as a factor in a 2 (personal, impersonal) x 2 (appropriate, inappropriate) ANOVA (10 subjects were excluded because of empty cells). RTs for "appropriate" and "inappropriate" responses were equivalent (F < 1), and response did not interact with the personal-impersonal variable, $F(1, 102) = 1.07$, $p_{rep} = .77$ (see Table 2). Thus, we failed to replicate the central behavioral result of Greene et al. (2001).

To examine whether RTs were affected by interactions between our new variables and response, we first conducted a 2 (self, other) x 2 (appropriate, inappropriate) ANOVA, which yielded only a significant interaction, $F(1, 97) = 6.97$, $\eta_p^2 = .07$, $p_{rep} > .99$ (15 subjects were excluded because of empty cells). When subjects were at risk, "appropriate" responses were significantly faster than "inappropriate" responses, $t(98) = -2.20$, $p_{rep} = .94$ (see Table 2). However, when only other people's lives were at risk, subjects were slower to endorse killing than they were to denounce it, $t(108) = 1.97$, $p_{rep} = .92$. Thus, subjects responded most quickly to

save themselves and others, as well as to abandon only others; in contrast, sacrificing themselves and saving only others required longer consideration. Regarding the inevitability factor, only the 2 (inevitable, avoidable) x 2 (appropriate, inappropriate) interaction yielded a significant effect, $F(1, 101) = 6.95$, $\eta_p^2 = .064$, $p_{rep} > .99$ (11 subjects were excluded because of empty cells). In avoidable dilemmas, subjects were slower to endorse killing than they were to denounce it (see Table 2), $t(106) = 1.73$, $p_{rep} = .89$; this finding suggests some deliberation over whether to take a life that would otherwise be spared. There was a weak effect in the opposite direction for inevitable dilemmas (see Table 2), $t(104) = 1.34$, $p_{rep} = .83$.

TABLE 2

Mean Response Times (in Milliseconds) by Dilemma Type and Response

	Response	
Dilemma type	"Appropriate"	"Inappropriate"
Personal	4,987 (2,611)	4,817 (2,503)
Impersonal	5,520 (2,912)	5,794 (3,117)
Self	4,988 (2,417)	5,706 (3,508)
Other	5,512 (2,697)	4,983 (2,291)
Inevitable	5,021 (2,225)	5,696 (3,191)
Avoidable	5,374 (3,102)	4,797 (2,342)
Instrumental	5,398 (3,155)	5,569 (3,358)
Incidental	6,227 (4,756)	6,015 (3,060)

Note. Values in parentheses are standard deviations of the means. All values were calculated from only those subjects who were included in the response time analyses.

Finally, the dual-process model of Greene et al. (2004) predicts that subjects with higher WMC will be faster to endorse personal killing than will subjects with lower WMC, because subjects with higher WMC better control Stroop-like response conflicts (e.g., Kane & Engle, 2003). A 2 (personal, impersonal) x 2 (appropriate, inappropriate) ANCOVA with WMC as the covariate found no support for the predicted three-way interaction (or for a main effect of WMC; both $Fs < 1$; 6 subjects were excluded because of empty cells). Indeed, in direct contradiction to the dual-process model of Greene et al. (2004), greater WMC predicted slightly longer RTs for "appropriate" responses to personal dilemmas ($r = .18$, $p_{rep} = .90$; all other $rs < .10$, $p_{rep}s < .75$).

DISCUSSION

Hypothetical moral judgments about killing one person to save others are affected by the personal or impersonal nature of inflicted harm, the benefit to the agent, the inevitability of victims' deaths, and individual differences in WMC. Our subjects found personal, direct harm to be especially inappropriate when it saved only other people, rather than themselves and other people, and when the harmed parties would not otherwise be killed. In contrast, they found impersonal, indirect harm to be particularly appropriate when it would save themselves and when the harmed parties would be killed regardless. Moreover, individual differences in WMC scores, which reflect cognitive-control variation, predicted moral judgments when personal killing harmed those whose fate was already sealed, as well as the consistency with which people made moral judgments about personal harm. These findings reinforce the importance and domain generality of WMC beyond typical laboratory tasks, demonstrate individual differences in moral judgment based on executive-control capability (despite noteworthy failures to identify individual differences in these judgments; Hauser et al., 2007), and clarify the particular roles of executive control in moral judgment. Although we replicated the effect of the personal-impersonal distinction obtained by Greene et al. (2001, 2004) and narrowed the definition of this distinction to relative directness, our findings do not support the idea that executive control must override prepotent emotional processes if people are to endorse provocative resolutions to moral dilemmas. First, WMC did not interact with the personal-impersonal variable alone in predicting judgments, as one would expect if cognitive control simply restrains emotional reactions. Subjects with higher WMC scores were more likely to endorse personal killing

only when harm was inevitable, and this increased endorsement would seem to result from deliberative reasoning, rather than simply from executive control over emotion. The association between higher WMC and longer RTs when responding "appropriate" to personal killing supports this interpretation, as does the greater consistency in judgments of personal killing among subjects with higher WMC scores (consistency across similar cases is a hallmark of rationality or reasoning; e.g., von Neumann & Morgenstern, 1947). Second, we did not replicate the Stroop-like finding (Greene et al., 2001) of slow affirmative responses to personal dilemmas. Instead, our subjects, with our improved stimulus materials, responded faster to personal dilemmas than to impersonal dilemmas, regardless of judgment.

Our WMC results may reflect a selectively engaged, voluntary reasoning system that can utilize WMC in the course of its functioning. For this suggestion to be correct, WMC must not predict judgments as a function of self-threat, if in fact the self-other distinction is processed automatically. We created the self-other variable with exactly this hypothesis in mind, and although further research is required to definitively determine whether the self-other distinction is processed intuitively, our data provide preliminary support: For self dilemmas, "appropriate" responses were faster and more frequent than "inappropriate" responses. The reverse pattern of responses and RTs was observed for other (i.e., non-self) dilemmas. The fact that the inevitable-avoidable variable did not exhibit this pattern and, unlike the self-other variable, did interact with WMC suggests that inevitability is processed with more effort and deliberation. Our findings are important because, if the self-other distinction is processed automatically, they suggest direct cooperation—outside of (or subsequent to) executive override functions—between some types of automatic-emotional processing and deliberative reasoning in generating moral judgments: The self-other and inevitable-avoidable variables had interactive effects only for impersonal dilemmas. A simple egoistic bias alone cannot explain why self-preservation and inevitability combined to increase "appropriate" responses only when subjects considered impersonal harm. Thus, automatically and more deliberatively processed variables may influence the formation of moral judgments when they converge, not only when they conflict (cf. Greene et al., 2004).

Notes:

[1]Many moral-judgment studies rely exclusively on subtle variations of the trolley problem (e.g., Hauser et al., 2007; Petrinovich, O'Neill, & Jorgensen, 1993; Valdesolo & DeSteno, 2006).

[2]One dilemma ("Modified Fumes") was miscoded and deleted prior to analyses.

[3]Data from 3 outliers were deleted. Response ("appropriate" vs. "inappropriate") was not included in this omnibus ANOVA because most subjects had no observations in at least one cell.

REFERENCES

Aquinas, T. (1988). On law, morality, and politics (W.P. Baumgarth, Ed., & R.J. Regan, Ed. & Trans.). Indianapolis, IN: Hackett.

Blair, R.J.R. (1995). A cognitive developmental approach to morality: Investigating the psychopath. Cognition, 57, 1–29.

Borg, J.S., Hynes, C., Van Horn, J., Grafton, S., & Sinnott-Armstrong, W. (2006). Consequences, action, and intention as factors in moral judgments: An fMRI investigation. Journal of Cognitive Neuroscience, 18, 803–817.

Braver, T.S., Gray, J.R., & Burgess, G.C. (2007). Explaining the many varieties of working memory variation: Dual mechanisms of cognitive control. In A.R.A. Conway, C. Jarrold, M.J. Kane, A. Miyake, & J.N. Towse (Eds.), Variation in working memory (pp. 76–106). Oxford, England: Oxford University Press.

Conway, A.R.A., Kane, M.J., Bunting, M.F., Hambrick, D.Z., Wilhelm, O., & Engle, R.W. (2005). Working memory span tasks: A methodological review and user's guide. Psychonomic Bulletin & Review, 12,769–786.

Cushman, F., Young, L., & Hauser, M. (2006). The role of conscious reasoning and intuition in moral judgment. Psychological Science, 17,1082–1089.

Engle, R.W., & Kane, M.J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. Ross (Ed.), The psychology of learning and motivation (pp. 145–199). New York: Academic Press.
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). An fMRI investigation of emotional engagement in moral judgment. Science, 293, 2105–2108.
Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. Psychological Review, 108,814–834.
Hasher, L., Lustig, C., & Zacks, R.T. (2007). Inhibitory mechanisms and the control of attention. In A.R.A. Conway, C. Jarrold, M.J. Kane, A. Miyake, & J.N. Towse (Eds.), Variation in working memory (pp. 227–249). Oxford, England: Oxford University Press.
Hauser, M., Cushman, F., Young, L., Jin, R., & Mikhail, J. (2007). A dissociation between moral judgments and justifications. Mind & Language, 22, 1–21.
Hinson, J.M., Jameson, T.L., & Whitney, P. (2003). Impulsive decision making and working memory. Journal of Experimental Psychology: Learning, Memory, and Cognition, 29, 298–306.
Kane, M.J., Bleckley, M.K., Conway, A.R.A., & Engle, R.W. (2001). A controlled-attention view of working-memory capacity. Journal of Experimental Psychology: General, 130, 169–183.
Kane, M.J., & Engle, R.W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. Journal of Experimental Psychology: General, 132, 47–70.
Kane, M.J., Hambrick, D.Z., & Conway, A.R.A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). Psychological Bulletin, 131, 66–71.
Killeen, P.R. (2005). Replicability, confidence, and priors. Psychological Science, 16, 1009–1012.
Lakoff, G. (2002). Moral politics: How liberals and conservatives think (2nd ed.). Chicago: University of Chicago Press.
Mikhail, J. (2007). Universal moral grammar: Theory, evidence, and the future. Trends in Cognitive Sciences, 11, 143–152.
Nichols, S. (2002). Norms with feeling: Towards a psychological account of moral judgment. Cognition, 84, 221–236.
Oberauer, K. (2005). Binding and inhibition in working memory: Individual and age differences in short-term recognition. Journal of Experimental Psychology: General, 134, 368–387.
Oberauer, K., Schulze, R., Wilhelm, O., & Süg, H.-M. (2005). Working memory and intelligence—their correlation and their relation: Comment on Ackerman, Beier, and Boyle (2005). Psychological Bulletin, 131, 61–65.
O'Neill, P., & Petrinovich, L. (1998). A preliminary cross-cultural study of moral intuitions. Evolution and Human Behavior, 19, 349–367.
Petrinovich, L., O'Neill, P., & Jorgensen, M. (1993). An empirical study of moral intuitions: Toward an evolutionary ethics. Journal of Personality and Social Psychology, 64, 467–478.
Pizarro, D.A., & Bloom, P. (2003). The intelligence of the moral intuitions: Comment on Haidt (2001). Psychological Review, 110, 193–196.
Prinz, J. (2006). The emotional basis of moral judgments. Philosophical Explorations, 9, 29–43.
Royzman, E.B., & Baron, J. (2002). The preference for indirect harm. Social Justice Research, 15, 165–184.
Spranca, M., Minsk, E., & Baron, J. (1991). Omission and commission in judgment and choice. Journal of Experimental Social Psychology, 27, 76–105.
Unsworth, N., Heitz, R.C., Schrock, J.C., & Engle, R.W. (2005). An automated version of the operation span task. Behavior Research Methods, 37, 498–505.
Valdesolo, P., & DeSteno, D. (2006). Manipulations of emotional context shape moral judgment. Psychological Science, 17, 476– 477.
von Neumann, J., & Morgenstern, O. (1947). Theory of games and economic behavior (2nd ed.). Princeton, NJ: Princeton University Press.