CAN Algorithm: An Individual Level Approach to identify Consequences and Norms Sensitivities and Overall Action/inaction Preferences in Moral Decision-making

Authors: Chuanjun Liu, Jiangqun Liao, Jiangqun Liao

Date: 2020-04-03T00:00:00+00:00

Abstract

Gawronski et al. (2017) developed a CNI model to measure an agent' s norms sensitivity, consequences sensitivity, and generalized inaction/action preferences when making moral decisions. However, the CNI model presupposed that an agent considers consequences–norms–generalized inaction/action preferences sequentially, which is untenable based on recent evidence. Moreover, the CNI model generates parameters at the group level based on binary categoric data. Hence, the C/N/I parameters cannot be used for correlation analyses or other conventional research designs. To solve these limitations, we developed the CAN algorithm to compute norms and consequences sensitivities and overall action/inaction preferences algebraically in a parallel manner. We re-analyzed the raw data of Gawronski et al.(2017) to test the methodological predictions. Our results demonstrate that: (1) the C parameter is approximately equal between the CNI model and CAN algorithm; (2) the N parameter under the CNI model approximately equals N/(1 -C) under the CAN algorithm; (3) the I parameter and A parameter are reversed around 0.5 -the larger the I parameter, the more the generalized inaction versus action preference and the larger the A parameter, the more overall action versus inaction preference; (4) tests of differences in parameters between groups with the CNI model and CAN algorithm led to almost the same statistical conclusion; (5) Parameters from the CAN algorithm can be used for correlational analyses and multiple comparisons, and this is an advantage over the parameters from the CNI model. The theoretical and methodological implications of our study were also discussed.

Full Text

Preamble

CAN Algorithm: An Individual-Level Approach to Identify Consequences and Norms Sensitivities and Overall Action/Inaction Preferences in Moral Decision-Making

Chuanjun Liu and Jiangqun Liao
Tsinghua University

Author Note

Correspondence concerning this article should be addressed to Jiangqun Liao, Department of Psychology, School of Social Sciences, Tsinghua University, Haidian District, Beijing, China, 100084. E-mail: liaojq@tsinghua.edu.cn. Tel & Fax: (86-10) 62787208.

Funding

This work was supported by the National Social Science Foundation of China [Grant No. 18BSH114] and the Tsinghua University Initiative Scientific Research Program [Grant No. 2017THZWYY11].

Abstract

Gawronski et al. (2017) developed a CNI model to measure an agent' s norms sensitivity, consequences sensitivity, and generalized inaction/action preferences when making moral decisions. However, the CNI model presupposes that an agent considers consequences–norms–generalized inaction/action preferences sequentially, which is untenable based on recent evidence. Moreover, the CNI model generates parameters at the group level based on binary categorical data, meaning that the C/N/I parameters cannot be used for correlation analyses or other conventional research designs. To address these limitations, we developed the CAN algorithm to compute norms and consequences sensitivities and overall action/inaction preferences algebraically in a parallel manner. We re-analyzed the raw data from Gawronski et al. (2017) to test our methodological predictions. Our results demonstrate that: (1) the C parameter is approximately equal between the CNI model and CAN algorithm; (2) the N parameter under the CNI model approximately equals N/(1 -C) under the CAN algorithm; (3) the I parameter and A parameter are reversed around 0.5–the larger the I parameter, the stronger the generalized inaction versus action preference, while the larger the A parameter, the stronger the overall action versus inaction preference; (4) tests of differences in parameters between groups with the CNI model and CAN algorithm led to almost identical statistical conclusions; and (5) parameters from the CAN algorithm can be used for correlational analyses and multiple comparisons, representing an advantage over parameters from the CNI model. The theoretical and methodological implications of our study are also discussed.

Keywords: CAN algorithm; moral dilemma; moral decision-making; CNI model

Introduction

Traditional moral dilemmas pit utilitarianism against deontology. Consider the well-known trolley problem: an uncontrollable trolley rushes toward five workers who do not notice the emergency. There is a sidetrack with only one worker, also unaware of the emergency. The only way to save the five workers is to pull a switch and divert the trolley onto the sidetrack, which would kill the one worker but save the five. The principle of utilitarianism is followed if the agent chooses to pull the switch because it achieves greater benefits than costs (Bentham, 1996; Mill, 1872). The principle of deontology is followed if the agent chooses not to pull the switch because harming the innocent violates moral norms (Kant & Gregor, 1997; Rawls, 1971).

However, interpretations of traditional moral dilemma paradigms are ambiguous (Gawronski, Armstrong, Conway, Friesdorf, & Hutter, 2017; Gawronski & Beer, 2017). In the trolley-car dilemma, there could be three reasons why an agent might pull the switch. First, the agent may have weaker norm sensitivity and be less averse to the sacrificial utilitarian proposal. Second, the agent may have stronger consequences sensitivity and find the outcome of pulling the switch considerably beneficial. Third, the agent may have a stronger generalized action preference (or weaker generalized inaction preference) irrespective of the norms and consequences involved. Traditional dilemma paradigms cannot dissociate these three possibilities, making it impossible to determine whether norm sensitivity, consequence sensitivity, or generalized action/inaction preferences drive the agent' s moral decision-making.

To resolve this ambiguity, Gawronski et al. (2017) developed a multinomial processing tree (MPT) model to dissociate these three interpretations. First, they expanded the conceptual manipulations of utilitarianism and deontology to address the limitations of traditional dilemmas. "Utilitarian" presupposes that observed behavior is sensitive to consequences, requiring experimental manipulations of consequences. "Deontological" presupposes that observed behavior is sensitive to moral norms, requiring experimental manipulations of moral norms.

Consequently, four types of dilemmas involving different combinations of consequences and norms must be considered (Gawronski et al., 2017; Gawronski & Beer, 2017): (a) a proscriptive norm opposes the proposed behavior, and the benefits for overall wellbeing are greater than the costs; (b) a proscriptive norm opposes the proposed behavior, and the benefits are smaller than the costs; (c) a prescriptive norm endorses the proposed behavior, and the benefits are greater than the costs; and (d) a prescriptive norm endorses the proposed behavior, and the benefits are smaller than the costs. Traditional moral dilemmas include only the first combined situation (proscriptive norm with benefits greater than costs) but not the other three.

Second, Gawronski et al. (2017) used an MPT to depict the mental processes underlying moral judgment [Figure 1: see original paper]. Together with the MPT model, the following equations predict action versus inaction responses. Since the sum of action and inaction probabilities in each dilemma equals 1, we list only the equations for action probability. To simplify notation, let p(action | proscriptive norm, benefits > costs) = p1, p(action | proscriptive norm, benefits < costs) = p2, p(action | prescriptive norm, benefits > costs) = p3, and p(action | prescriptive norm, benefits < costs) = p4.

$$p1 = C + (1 - C) \times (1 - N) \times (1 - I) \tag{1}$$
$$p2 = (1 - C) \times (1 - N) \times (1 - I) \tag{2}$$
$$p3 = C + (1 - C) \times N + (1 - C) \times (1 - N) \times (1 - I) \tag{3}$$
$$p4 = (1 - C) \times N + (1 - C) \times (1 - N) \times (1 - I) \tag{4}$$

This model can dissociate three parameters using maximum likelihood statistics: Consequences sensitivity (C), Norms sensitivity (N), and generalized Inaction versus Action irrespective of consequences and norms (I). Hence, the model was termed the "CNI model." Gawronski et al. (2017) provided protocols with the MultiTree program (Moshagen, 2010) to generate C/N/I parameters (www.bertramgawronski.com/documents/CNI-Model_{Materials}.zip).

Methodological Limitations of the CNI Model

The CNI model contributes to the literature by purportedly dissociating the three possibilities when agents make decisions in traditional moral dilemmas, thereby helping to resolve inconsistent findings, such as whether and how incidental emotions affect moral judgment (Gawronski, Conway, Armstrong, Friesdorf, & Hutter, 2018). However, recently Baron and Goodwin demonstrated several theoretical problems underlying the CNI model, such as the prohibition of deontological rules (for details, see Baron & Goodwin, 2019). In the present study, we highlight methodological limitations of the CNI model and propose a new algorithm to address them.

First, the CNI model cannot be applied to correlation and regression analyses. C/N/I parameters are generated using maximum likelihood statistics at the group level rather than the individual level, making the model unsuitable for studies examining correlations.

Second, the CNI model can only compare differences between two parameters or one parameter to a specific value. It is inapplicable for multiple comparisons beyond two conditions.

Most importantly, the CNI model hypothesizes that the agent sequentially considers whether the consequences of the proposed behavior are beneficial, then whether the behavior is permitted by moral norms, and finally considers generalized action or inaction irrespective of consequences or norms. This a priori hypothesis is untenable for two reasons. First, if agents sequentially considered decision principles, they would not experience dilemma when norms prohibit action while consequences advocate it. Agents feel dilemmatic only when simultaneously considering conflicting norms and consequences principles. Thus, agents are more likely to activate norms and consequences principles in parallel rather than sequentially.

Second, even if agents adopt a sequential mindset, other processing patterns are possible, such as N→C→I (first consider norms, then consequences, then generalized action/inaction), I→C→N (first obtain a generalized action/inaction preference, then revise by consequences, then by norms), I→N→C (first obtain a generalized action/inaction preference, then revise by norms, then by consequences), and other potential sequences. Taking the N→C→I pattern as an example (named the "NCI model"), see [Figure 2: see original paper].

With the NCI model, we can also depict response probabilities for the four combined dilemma situations:

$$p1 = (1 - N) \times C + (1 - N) \times (1 - C) \times (1 - I) \tag{5}$$
$$p2 = (1 - N) \times (1 - C) \times (1 - I) \tag{6}$$
$$p3 = N + (1 - N) \times C + (1 - N) \times (1 - C) \times (1 - I) \tag{7}$$
$$p4 = N + (1 - N) \times (1 - C) \times (1 - I) \tag{8}$$

Gawronski et al. (2017) noted in their footnote 7 that all reported effects were replicated with the NCI model, with only minor differences where some marginally significant effects in the CNI model became statistically significant with the NCI model. Therefore, they did not further discuss differences between the models. However, if the CNI and NCI models depict observed data equally well, equations (1)-(4) and equations (5)-(8) would be statistically identical for parameter generation. Taking the N parameter as an example, it can be transformed from equations (1)-(4) to yield N = (−p1 − p2 + p3 + p4)/(2 − p1 + p2 − p3 + p4), and from equations (5)-(8) to yield N = (−p1 − p2 + p3 + p4)/2. If the models were statistically equivalent, these two N parameters should be equal. After conversion, this implies p2 − p1 = p3 − p4. Similarly, transforming the C parameter from both models yields p1 + p2 = p3 + p4. Combining these equations implies p2 = p3 and p1 = p4. Thus, the CNI and NCI models would generate identical N and C parameters only if p2 = p3 and p1 = p4, a precondition with very low empirical likelihood.

The first two limitations stem from parameters being recorded at the group level rather than individual level. The most serious limitation arises because the CNI model presupposes sequential rather than parallel consideration of norms and consequences principles. Given these methodological limitations, we developed a new algorithm to identify agents' norms and consequences sensitivities and overall action/inaction preferences.

The CAN Algorithm

The traditional moral dilemma is varied into four parallel versions by manipulating the potential moral principles of norms and consequences, where action is either prohibited or advocated by these principles. We can therefore use a common algebraic subtraction strategy to generate C and N parameters, similar to approaches used in the literature (e.g., Talhelm et al., 2014 computed loyalty/nepotism as the amount participants rewarded their friend minus the amount they punished their friend). For the A parameter, we used an aggregate mean strategy to measure overall action versus inaction preferences.

For the C parameter, if individuals are sensitive to consequences, they should be more likely to approve proposals when benefits exceed costs than when benefits are smaller than costs. Therefore, consequences sensitivity under proscriptive norms can be represented by p1 -p2, and under prescriptive norms by p3 -p4. Hence, consequences sensitivity is the mean of these two conditions: C = (p1 - p2 + p3 -p4)/2.

For the N parameter, norms sensitivity under conditions where benefits exceed costs can be represented by p3 -p1, and where benefits are smaller than costs by p4 -p2. Thus, norms sensitivity is the mean of these two conditions: N = (p3 -p1 + p4 -p2)/2.

For the A parameter, this index represents an individual's overall action/inaction preferences rather than generalized preferences irrespective of norms and consequences. The mean action probability across the four situations is calculated as A = (p1 + p2 + p3 + p4)/4. We argue that the I parameter under sequential processing in the CNI model lacks theoretical coherence, as discussed below.

If the C/N parameter is greater (or less) than 0, participants are identified as sensitive to supporting (or opposing) norms/consequences. Larger C/N parameters indicate greater sensitivity. If the C/N parameter does not differ significantly from 0, participants are identified as insensitive to norms/consequences. Larger A parameters indicate greater overall endorsement of the behavioral proposal. If the A parameter is greater (or less) than 0.5, participants are identified as having an overall action (or inaction) preference. If the A parameter does not differ significantly from 0.5 while at least one C/N parameter differs significantly from 0, participants are identified as having a pure moral attitude (utilitarian or deontological). If the A parameter does not differ significantly from 0.5 and neither C nor N parameters differ significantly from 0, participants are identified as responding randomly.

To differentiate it from the CNI model, we named this new algorithm "CAN." To further demonstrate its reasonableness and differences from the CNI model, we discuss the equations of both approaches.

ChinaRxiv 中X

Contrasts in Parameters Between the CNI Model and CAN Algorithm

The equations for the C parameter are identical under both methods. Using the CNI model equations, we can transform equations (1) and (2) to yield C = p1 -p2, and equations (3) and (4) to yield C = p3 -p4. On average, C = (p1 -p2 + p3 -p4)/2, which is identical to the CAN algorithm equation. Therefore, we predict that in each study from Gawronski et al. (2017), the mean C parameter under the CNI model will be almost equal to the mean C parameter under the CAN algorithm. Given that the CNI model' s C parameter was computed with maximum likelihood statistics at the group level while the CAN algorithm' s C parameter was computed algebraically at the individual level, these values should be approximately rather than absolutely equal.

For the N parameter, we can transform equations (1)-(4) of the CNI model to yield N = (p3 -p1 + p4 -p2)/(2 × (1 -C)). However, the CAN algorithm' s N parameter equation is N = (p3 -p1 + p4 -p2)/2. Thus, the CAN algorithm' s N parameter divided by (1 -C) should approximate the CNI model' s N parameter. This occurs because the CNI model hypothesizes that agents consider norms only after not considering consequences, an untenable precondition since agents may consider norms first, suggesting the NCI model as the appropriate sequential alternative. If so, transforming equations (5)-(8) of the NCI model yields N = (p3 -p1 + p4 -p2)/2, identical to the CAN algorithm. Thus, we predict that N/(1 -C) under the CAN algorithm will approximate the N parameter value under the CNI model.

Regarding I and A parameters, the CNI model logic suggests that agents consider generalized action/inaction preferences only after not considering norms and consequences. Similar to the N parameter, this precondition is unreasonable. Agents may first develop a generalized action/inaction preference, which is then influenced by norms and consequences principles. Therefore, the CNI model' s I parameter lacks credibility. We abandoned attempts to identify generalized inaction/action preferences irrespective of norms and consequences, instead focusing on the overall action/inaction tendency as more critical. If agents decide purely based on norms and consequences principles, p2 would approach 0 and p3 would approach 1 (since norms and consequences prohibit or advocate action), while the mean of p1 and p4 would approach 0.5 (since norms and consequences conflict). Thus, the A parameter ((p1 + p2 + p3 + p4)/4) should not differ from 0.5 if agents decide purely based on norms and consequences or respond randomly. Consequently, the A parameter can represent agents' overall action/inaction preferences.

Overview of the Present Study

Based on the methodological discussion above, we formulated several predictions. The C parameter under the CNI model will approximate the C parameter under the CAN algorithm (H1). The N parameter under the CNI model will approximate N/(1 -C) under the CAN algorithm (H2). The I parameter under the CNI model represents generalized inaction versus action preferences, whereas the A parameter under the CAN algorithm represents overall action versus inaction preferences. Thus, these parameters will be reversed around 0.5: if the I parameter exceeds 0.5, the A parameter will be lower than 0.5, and vice versa (H3). Although the CNI model and CAN algorithm are algebraically different, biases may be systematically balanced across between-subject conditions, leading to almost identical tests of between-subject differences in parameters from both approaches (H4).

Methods

To test these predictions, we re-analyzed the raw data from Gawronski et al. (2017), where the CNI model was originally proposed and tested. We first downloaded the raw data from https://osf.io/xt66w/. We then re-analyzed the data with the CNI model to ensure reproducibility of Gawronski et al.' s (2017) results, which were identical to their reported findings. Second, we applied the CAN algorithm to generate C/N/A parameters and calculated N/(1 -C) using mean C and N values. Afterward, we tested the hypotheses stated above. Finally, because CAN algorithm parameters are at the individual level, they can be used for correlation and other analyses. We therefore conducted Pearson correlation analyses between psychopathy scale ratings and parameters from Gawronski et al.' s (2017) Study 4.

Results

Gawronski et al. (2017) conducted four formal studies and one supplemental study, each replicated based on concerns about reproducibility in psychological research (Open Science Collaboration, 2015), yielding 10 studies total. Study 1a/b examined gender differences in moral decision-making, which remained ambiguous in traditional dilemma approaches (Friesdorf, Conway, & Gawronski, 2015). Study 2a/b explored effects of cognitive load on moral decision-making (Greene, Morelli, Lowenberg, Nystrom, & Cohen, 2008). Study 3a/b investigated question-framing effects, as prior work showed personal force enhanced deontological responses (Greene, Sommerville, Nystrom, Darley, & Cohen, 2001). Study 4a/b examined relationships between subclinical psychopathy levels and utilitarian responses (Bartels & Pizarro, 2011; Kahane, Everett, Earp, Farias, & Savulescu, 2015). Study S1a/b investigated effects of harm salience on moral decision-making (Conway & Gawronski, 2013).

Gawronski et al. (2017) analyzed traditional dilemmas, process dissociation (Conway & Gawronski, 2013), and the CNI model for each study. In our re-analysis, we focused exclusively on the CNI model and re-analyzed raw data using the CAN algorithm. The re-analysis patterns were almost identical across all studies. We present results for Studies 1a/1b here; remaining results appear in the Appendices.

Hypothesis Testing

Re-analysis of Gawronski et al. (2017) Study 1a. [Figure 3: see original paper] displays results from the CNI model (left) and CAN algorithm (right). Error bars represent $±1*SE.presentstestsofgenderdifferencesusingbothapproaches.TheCparametershowedno*
$\Delta G^{\{2\}(1) = 1.34, p = .247; CAN: t(199) = 1.08, p = .281).TheNparametershowedsignificantgenderdifferencesunderbothmodels(CNI}$
$26.00, p < .001; CAN: t(199) = 2.74, p = .007).TheIparameterundertheCNImodelandAparameterundertheC$
$\Delta G^{2}(1) = 12.34, p < .001; CAN: t(199) = 2.45, p = .015), consistent with their reversed interpretations around 0.5. N/(1 -C) under the CAN algorithm was approximately equal to the N parameter under the CNI model.

Re-analysis of Gawronski et al. (2017) Study 1b. [Figure 4: see original paper] shows results from the CNI model (left) and CAN algorithm (right). Error bars represent $±1*SE.presentsgenderdifferencetests.TheCparametershowedsignificantgenderdiffere*
$\Delta G^{\{2\}(1) = 6.43, p = .011, d = 0.364; CAN: t(195) = 2.39, p = .018).TheNparameteralsoshowedsignificantgenderdifferences(CNI: \Delta}$
$17.43, p < .001, d = 0.599; CAN: t(195) = 2.20, p = .029).TheIandAparametersagainshowedreversedpattern$
$\Delta G^{2}(1) = 9.12, p = .003, d = 0.428; CAN: t(195) = 2.41, p = .017). N/(1 -C) approximated the CNI N parameter.

Across all 10 studies, predictions were validated. The C parameter was approximately equal between models. The N parameter from the CAN algorithm was slightly smaller than from the CNI model, with N/(1 -C) approximating the CNI N parameter. The I and A parameters were reversed around 0.5, as their statistical implications differ: larger I indicates more generalized inaction versus action preference, while larger A indicates more overall action versus inaction tendency. Furthermore, between-subject parameter differences were almost identical across models, though independent-sample t-tests with C/A/N parameters were more stringent than maximum likelihood statistics with C/N/I parameters, causing some marginally significant CNI results to become non-significant under the CAN algorithm (see Appendices).

Statistical Advantage of the CAN Algorithm

The CAN algorithm enables correlation analysis, demonstrating a statistical advantage over the CNI model. Using Study 4 data, where participants were artificially divided into "low" and "high" psychopathy conditions based on scale scores, we conducted Pearson correlations between psychopathy scores and C/A/N parameters. In Study 4a, psychopathy scores did not correlate significantly with the C parameter (r = −0.122, p = .100) but correlated significantly with the N parameter (r = −0.153, p = .038) and not with the A parameter (r = 0.106, p = .152). These results align with CAN algorithm difference tests but not with C/I parameter difference tests (see Appendix Table A5). In Study 4b, psychopathy correlated significantly with both C (r = −0.307, p < .001) and N parameters (r = −0.394, p < .001) but not with the A parameter (r = 0.103, p = .149). These results also support CAN algorithm difference tests but not I parameter tests (Appendix Table A6). Overall, correlation analyses supported CAN algorithm results while varying slightly from CNI model results. Methodologically, the

CAN algorithm generates individual-level parameters suitable for correlation and other common analyses, representing an advantage over the CNI model.

Discussion

All hypotheses were verified, and additional correlation analyses demonstrated that CAN algorithm parameters can be used for correlational analyses—an advantage over the CNI model. Because C/A/N parameters are grounded at the individual level, they support a wider range of conventional analyses across research designs.

Conceptual Manipulation Development

The CAN algorithm builds on the CNI model' s theoretical foundation. Originally, traditional dilemmas considered only scenarios where proscriptive norms coincided with benefits greater than costs (e.g., trolley-car paradigm; Foot, 1967; footbridge paradigm; Thomson, 1976). Conway and Gawronski (2013) later explored two scenario types: consistent edition (proscriptive norms with benefits smaller than costs) and inconsistent edition (proscriptive norms with benefits greater than costs). Gawronski et al. (2017) varied both norms and consequences, creating four scenario types: proscriptive/prescriptive norms with benefits greater/smaller than costs. This conceptual manipulation deepened insights into moral decision-making.

However, criticisms remain. Baron and Goodwin (2019) questioned whether the norms underlying proscriptive and prescriptive editions represent the same moral principle. For example, in a transplant dilemma, proscriptive editions invoke the norm against harming others, while prescriptive editions invoke the norm of stopping harmful actions—essentially different moral norms.

Nevertheless, we believe the four-edition manipulation contributes meaningfully. It is event-oriented rather than principle-oriented: across the four editions, the event remains consistent while underlying norms and consequences vary. Whether agents are sensitive to changes in norms and consequences for the same event is important. Moreover, proscriptive and prescriptive morality are crucial in daily life as two facets of moral regulation (Janoff-Bulman, Sheikh, & Hepp, 2009). While we acknowledge that proscriptive and prescriptive scenarios involve different norms, we do not believe this undermines norm sensitivity measurement.

Methodological Development of the CAN Algorithm

The CAN algorithm preserves the theoretical conceptual manipulation while addressing CNI model limitations. The CNI model cannot be used for correlation analysis or multiple comparisons because its parameters are group-level rather than individual-level. The CAN algorithm generates parameters algebraically at the individual level and can be applied to continuous-scale data, not just binary categorical data. Therefore, C/A/N parameters support a broader scope of research designs and statistical analyses.

The most serious methodological limitation is the CNI model' s presupposition that agents sequentially consider consequences, norms, and generalized inaction/action preferences. This questionable precondition artificially overestimates the N parameter. As our re-analysis demonstrated, the CNI model' s N parameter approximately equals N/(1 -C) from the CAN algorithm. Since (1 -C) ranges from [0, 1] in the CAN algorithm, the CNI model' s N parameter is systematically larger than the CAN algorithm' s N parameter. The sequential process assumption also renders the I parameter dubious, as it supposedly depicts generalized inaction/action preferences independent of norms and consequences. Therefore, the CAN algorithm adopts a standard subtraction strategy for C and N parameters and establishes an overall action/inaction preference index: the A parameter.

Even if moral decision-making occurs sequentially, the NCI model is more credible than the CNI model. Theoretical models of social intuition suggest that people react emotionally first and revise cognitively later, considering norms intuitively before consequences rationally (Haidt, 2001). Thus, the NCI model is more reasonable even under sequential processing assumptions. However, increasing evidence suggests that emotional and cognitive processes are parallel and independent rather than sequential (Cushman, Young, & Greene, 2010; Greene, 2009; Hutcherson, Montaser-Kouhsari, Woodward, & Rangel, 2015; Paxton & Greene, 2010). Therefore, the CAN algorithm more appropriately demonstrates moral preferences, particularly for C and N parameters.

A More Reasonable Multinomial Processing Tree Model Supports the CAN Algorithm

Processing tree models provide powerful frameworks for examining potentially conflicted cognitive processes (Calanchini, Rivers, Klauer, & Sherman, 2018; Hutter & Klauer, 2016). We therefore constructed an alternative MPT based on theoretical and empirical evidence [Figure 5: see original paper], named the "DNA model." Bago and De Neys (2019) proposed a corrective dual-process model of moral cognition, finding that participants were intuitively utilitarian in a two-response paradigm and concluding that final moral judgments depend on the absolute and relative strength of competing deontological and utilitarian intuitions. In our MPT, we hypothesized that driving forces from moral norms and consequences operate in parallel, with response patterns determined by which force is stronger.

Together with the DNA model, four equations can be constructed:

$$p1 = (1–D) \times A + D \times (1–N) \tag{9}$$
$$p2 = (1–D) \times A \tag{10}$$
$$p3 = D \times N + (1–D) \times A + D \times (1–N) \tag{11}$$
$$p4 = D \times N + (1–D) \times A \tag{12}$$

In the DNA model, consequences sensitivity and norms sensitivity can also be calculated algebraically. Consequences sensitivity is represented by D × (1 – N) because agents must first be sensitive to moral principles and then to the consequences principle. This algebraic expression equals (p1 -p2 + p3 -p4)/2 based on equations (9)-(12). Similarly, norms sensitivity is represented by D × N, which transforms to (p3 -p1 + p4 -p2)/2 from equations (9)-(12). These indices are identical to the CAN algorithm. Indeed, the DNA model more accurately depicts moral decision-making processes according to the literature (Bago & De Neys, 2019; Neys & Pennycook, 2019).

Re-examining the CNI model's I parameter from the DNA model perspective reveals its problems. In the DNA model, general preference for action irrespective of moral principles is ((1 -D) × A), with probability p2 (when agents endorse behavior prohibited by both norms and consequences). General preference for inaction irrespective of moral principles is ((1 -D) × (1 -A)), with probability (1 -p3) (when agents decline behavior advocated by both norms and consequences). Thus, in the DNA model, generalized action preference irrespective of norms and consequences is p2, while generalized inaction preference is 1 -p3. Gawronski et al. (2017) hypothesized that the sum of these preference probabilities equals 1. If so, p2 + (1 -p3) = 1, implying p2 = p3. This means the sum of generalized action and inaction preference probabilities could equal 1 only when p2 = p3, which is highly improbable. Therefore, the CNI model' s I parameter, purporting to depict generalized inaction/action preferences independent of norms and consequences, lacks credibility. In contrast, the CAN algorithm' s A parameter is methodologically credible as it depicts overall action/inaction preferences across the four dilemma editions.

Conclusion

In summary, we addressed methodological limitations of the CNI model and resolved them with a new algorithm: CAN. The CNI model presupposes that agents sequentially consider consequences, norms, and generalized inaction/action preferences in moral decision-making. We provided theoretical evidence that decision-making is more likely parallel for norms and consequences and developed the CAN algorithm accordingly. While the CNI model generates group-level parameters, we calculated parameters algebraically at the individual level, making the CAN algorithm suitable for a broader range of research designs and conventional statistical analyses.

References

Bago, B., & De Neys, W. (2019). The intuitive greater good: Testing the corrective dual process model of moral cognition. *Journal of Experimental Psychology: General, 148*(10), 1782-1795.

Baron, J., & Goodwin, G. P. (2019). Consequences, Norms, and Inaction: A Comment. *SSRN Electronic Journal*. doi:10.2139/ssrn.3378106

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.

Bentham, J. (1996). *An Introduction to the Principles of Morals and Legislation*. New York: Oxford University Press, USA. (Original work published 1781).

Calanchini, J., Rivers, A. M., Klauer, K. C., & Sherman, J. W. (2018). Chapter Two - Multinomial processing trees as theoretical bridges between cognitive and social psychology. In K. D. Federmeier (Ed.), *Psychology of Learning and Motivation* (pp. 39-65): Academic Press.

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), 216-235.

Cushman, F., Young, L., & Greene, J. D. (2010). Multi-system moral psychology. In J. D. e. al. (Ed.), *The Moral Psychology Handbook* (pp. 47-71). New York: Oxford University Press.

Foot, P. (1967). The Problem of Abortion and the Doctrine of Double Effect. *Oxford Review, 2*(2), 152–161.

Friesdorf, R., Conway, P., & Gawronski, B. (2015). Gender differences in responses to moral dilemmas: a process dissociation analysis. *Personality and Social Psychology Bulletin, 41*(5), 696-713.

Gawronski, B., Armstrong, J., Conway, P., Friesdorf, R., & Hutter, M. (2017). Consequences, norms, and generalized inaction in moral dilemmas: The CNI model of moral decision-making. *Journal of Personality and Social Psychology, 113*(3), 343-376.

Gawronski, B., & Beer, J. S. (2017). What makes moral dilemma judgments "utilitarian" or "deontological" ? *Social Neuroscience, 12*(6), 626-632.

Gawronski, B., Conway, P., Armstrong, J., Friesdorf, R., & Hutter, M. (2018). Effects of incidental emotions on moral dilemma judgments: An analysis using the CNI model. *Emotion, 18*(7), 989-1008.

Greene, J. D. (2009). Dual-process morality and the personal/impersonal distinction: A reply to McGuire, Langdon, Coltheart, and Mackenzie. *Journal of Experimental Social Psychology, 45*(3), 581-584.

Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition, 107*(3), 1144-1154.

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.

Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological Review, 108*(4), 814-834.

Hutcherson, C. A., Montaser-Kouhsari, L., Woodward, J., & Rangel, A. (2015). Emotional and Utilitarian Appraisals of Moral Dilemmas Are Encoded in Separate Areas and Integrated in Ventromedial Prefrontal Cortex. *Journal of Neuroscience, 35*(36), 12593-12601.

Hutter, M., & Klauer, K. C. (2016). Applying processing trees in social psychology. *European Review of Social Psychology, 27*(1), 116-159.

Janoff-Bulman, R., Sheikh, S., & Hepp, S. (2009). Proscriptive versus prescriptive morality: two faces of moral regulation. *Journal of Personality and Social Psychology, 96*(3), 521-537.

Kahane, G., Everett, J. A., Earp, B. D., Farias, M., & Savulescu, J. (2015). 'Utilitarian' judgments in sacrificial moral dilemmas do not reflect impartial concern for the greater good. *Cognition, 134*, 193-209.

Kant, I., & Gregor, M. J. (1997). *Groundwork of the metaphysics of morals*. Cambridge, UK: Cambridge University Press.

Mill, J. S. (1872). *The Logic of the Moral Sciences*. Chicago, Illinois, US: Open Court.

Moshagen, M. (2010). multiTree: a computer program for the analysis of multinomial processing tree models. *Behavior Research Methods, 42*(1), 42-54.

Neys, W. D., & Pennycook, G. (2019). Logic, Fast and Slow: Advances in Dual-Process Theorizing. *Current Directions in Psychological Science, 28*(5), 503-509.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251), aac4716.

Paxton, J. M., & Greene, J. D. (2010). Moral reasoning: hints and allegations. *Topics in Cognitive Science, 2*(3), 511-527.

Rawls, J. R. (1971). *A Theory of Justice*. Cambridge, Massachusetts, USA; London, England: The Belknap Press of Harvard University Press.

Talhelm, T., Zhang, X., Oishi, S., Shimin, C., Duan, D., Lan, X., & Kitayama, S. (2014). Large-Scale Psychological Differences Within China Explained by Rice Versus Wheat Agriculture. *Science, 344*(6184), 603-608.

Thomson, J. J. (1976). Killing, letting die, and the trolley problem. *Monist, 59*(2), 204-217.

Appendices

Re-analysis of Gawronski et al. (2017) Study 2a

[FIGURE:A1] displays results from the CNI model (left) and CAN algorithm (right) for Study 2a. Error bars represent $±1SE.[TABLE: A1]presentstestsofcognitiveloaddifferences.TheCparametershowednosignificantloadeffectundereitherm
ΔG{2}(1) = 1.35, p = .245, d = 0.168; CAN : t(192) = 1.22, p = .223).TheNparameteralsoshowednosignificanteffect(CNI : ΔG{2}(1)
0.01, p = .927, d = 0.013; CAN : t(192) = 0.22, p = .827).TheIandAparametersshowedsignificanteffects(CN
ΔG^{2}$(1) = 5.19, p = .023, d = 0.328; CAN: t(192) = 2.01, p = .045), with patterns reversed around 0.5 as predicted.

Re-analysis of Gawronski et al. (2017) Study 2b

[FIGURE:A2] shows Study 2b results. Error bars represent $±1SE.[TABLE: A2]presentscognitiveloaddifferencetests.TheCparametershowednosignificantloadeffect(CNI :
ΔG{2}(1) = 2.08, p = .149, d = 0.209; CAN : t(192) = 1.45, p = .149).TheNparameteralsoshowednoeffect(CNI : ΔG{2}(1) =
0.05, p = .826, d = 0.032; CAN : t(192) = 0.09, p = .927).TheIandAparametersshowedsignificanteffects(CN
ΔG^{2}$(1) = 13.77, p < .001, d = 0.535; CAN: t(192) = 2.97, p = .003), with reversed patterns around 0.5.

Re-analysis of Gawronski et al. (2017) Study 3a

[FIGURE:A3] presents Study 3a results. Error bars represent $±1SE.[TABLE: A3]showsquestion−framingdifferencetests.TheCparametershowednosignificantframingeffect(CNI :
ΔG{2}(1) = 2.44, p = .118, d = 0.230; CAN : t(184) = 1.39, p = .168).TheNparametershowedmarginallysignificanteffects(CNI : ΔG{2
3.31, p = .069, d = 0.268; CAN : t(184) = 1.47, p = .144).TheIandAparametersshowedsignificanteffects(CN
ΔG^{2}$(1) = 35.18, p < .001, d = 0.713; CAN: t(184) = 4.43, p < .001), with reversed patterns around 0.5.

Re-analysis of Gawronski et al. (2017) Study 3b

[FIGURE:A4] displays Study 3b results. Error bars represent $±1SE.[TABLE: A4]presentsframingdifferencetests.TheCparametershowednosignificanteffect(CNI :
ΔG{2}(1) = 0.09, p = .767, d = 0.043; CAN : t(187) = 0.36, p = .721).TheNparametershowedsignificanteffectsundertheCNImodel(ΔG
6.15, p = .013, d = 0.363)butnotundertheCANalgorithm(t(187) = 1.46, p = .147), representingadiscrepancy.TheIandAparametersshowedsignificanteffects(CNI :
ΔG^{2}$(1) = 29.50, p < .001, d = 0.799; CAN: t(187) = 4.62, p < .001), with reversed patterns around 0.5.

Re-analysis of Gawronski et al. (2017) Study 4a

[FIGURE:A5] shows Study 4a results. Error bars represent $±1SE.[TABLE: A5]presentspsychopathydifferencetests.TheCparametershowedmarginallysignificanteffects(CNI :

$\Delta G^{\{2\}(1) = 2.77, p = .096, d = 0.247; CAN: t(182) = 1.56, p = .121). The N parameter showed significant effects under the CNI model (\Delta G}$
$12.35, p < .001, d = 0.521) but marginal effects under the CAN algorithm (t(182) =$
$1.89, p = .060), representing a basic identity with slight discrepancy. The I and A parameters showed marginal eff$
$\Delta G^{2}(1) = 3.15, p = .076, d = 0.262; CAN: t(182) = 1.22, p = .223), with reversed patterns around 0.5.

Re-analysis of Gawronski et al. (2017) Study 4b

[FIGURE:A6] presents Study 4b results. Error bars represent $\pm 1 SE. [TABLE: A6] shows psychopathy difference tests. The C parameter showed significant effects under both models (CNI:$
$\Delta G^{\{2\}(1) = 23.13, p < .001, d = 0.695; CAN: t(196) = 4.40, p < .001). The N parameter also showed significant effects (CNI: \Delta G} \{2\}(1)$
$111.80, p < .001, d = 1.48; CAN: t(196) = 6.12, p < .001). The I and A parameters showed significant effects un$
= 8.90, p = .003, d = 0.406) but not under the CAN algorithm (t(196) = 1.28, p = .202), representing a discrepancy (note: Levene' s test was significant, p < 0.05, suggesting violation of equal variance assumption).

Re-analysis of Gawronski et al. (2017) Study S1a

[FIGURE:A7] displays Study S1a results. Error bars represent $\pm 1 SE. [TABLE: A7] presents harm salience difference tests. No parameters showed significant salience effects: C parameter (CNI: \Delta G^{\{2\}(1) = 2.15, p = .143, d = 0.211; CAN: t(193) = 1.42, p = .156), N parameter (CNI: \Delta G} \{2\}(1) =$
$0.10, p = .758, d = 0.044; CAN: t(193) = 0.39, p = .696), and I/A parameters (CNI:$
$\Delta G^{2}(1) = 1.60, p = .206, d = 0.182; CAN: t(193) = 0.97, p = .334).

Re-analysis of Gawronski et al. (2017) Study S1b

[FIGURE:A8] shows Study S1b results. Error bars represent $\pm 1 SE. [TABLE: A8] presents harm salience difference tests. The C parameter showed significant effects (CNI:$
$\Delta G^{\{2\}(1) = 4.40, p = .036, d = 0.305; CAN: t(189) = 2.20, p = .029). The N parameter showed significant effects (CNI: \Delta G} \{2\}(1) =$
$15.79, p < .001, d = 0.580; CAN: t(189) = 2.11, p = .036). The I and A parameters showed no significant effects$
$\Delta G^{2}(1) = 0.97, p = .325, d = 0.144; CAN: t(189) = 0.58, p = .563).

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv —Machine translation. Verify with original.