Evolutionary Behavioral Sciences
2018, Vol. 12, No. 3, 206–217

2330-2925/18/$12.00 http://dx.doi.org/10.1037/ebs0000114

Testing Error-Management Predictions in Forgiveness Decisions With Cognitive Modeling and Process-Tracing Tools

Jolene H. Tan, Shenghua Luan, Tita Gonzalez, and Evaldas Jablonskis
Max Planck Institute for Human Development, Berlin, Germany

We investigated the forgiveness decision as an error-management task and demonstrated how tools from decision science can facilitate testing precise predictions about bias and its cognitive implementation. We combined decision modeling (using a weighting-and-adding model and a lexicographic heuristic) with process-tracing tools that track response times as well as the pattern of information acquisition. Our modeling results indicate that individuals adopted a decision bias commensurate with the relative cost of errors and that they also adjusted their bias after the perceived costs of errors were experimentally manipulated. Even though the 2 decision models were accurate in fitting the decisions (accuracies of around 85%), they were less successful in fitting the process measures. Our process-tracing results do not support either model—response times were in favor of the heuristic, whereas information-acquisition patterns favored the linear model, albeit slightly. Nevertheless, our methodology used to investigate the forgiveness decision can be a seen as a "blueprint" of how the cognitive processes of other error-management tasks can be investigated and how a more detailed mapping of the adapted mind can be achieved.

Keywords: forgiveness, fast-and-frugal trees, Franklin's rule, signal-detection theory, error-management theory

Supplemental materials: http://dx.doi.org/10.1037/ebs0000114.supp

Engineering a smoke alarm involves considering how to balance two potential errors: Should it be calibrated so that even a wisp of smoke from candlelight will set it off? Or should it ring only if it detects a thick gulf of smoke? Because most people would agree that it is better for a smoke alarm to ring even when there is no danger than for it to stay silent during a fire, most smoke alarms are calibrated to be biased like the former rather than the latter. Error-management theory (EMT) and signal detection theory (SDT), the study of the mind based on this "smoke detector principle," have generated insights across multiple domains about how biases can be adaptive if they reduce the likelihood of the more costly error (e.g., Green & Swets, 1966; Haselton et al., 2009; Nesse, 2005).

In an earlier article (Tan, Luan, & Katsikopoulos, 2017), we proposed that deciding whether to forgive is an error-management task and showed how the decision process can be studied using cognitive modeling. In this article, we augment the modeling approach with process-tracing tools and experimental manipulation of bias. The goal of the current research was to extend that work by demonstrating how EMT can lead to predictions about the cognitive processes that implement bias. Here, we show how cognitive modeling and process-tracing tools can be combined to facilitate more precise tests about bias on an individual level as well as about whether biases are updated after error costs have changed.

This article was published Online First December 18, 2017.

Jolene H. Tan, Shenghua Luan, Tita Gonzalez, and Evaldas Jablonskis, Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, Berlin, Germany.

Tita Gonzalez is now at the Department of Psychology, University of Basel. Evaldas Jablonskis is now at the Department of Philosophy, University of Vienna.

We thank Amit Kothiyal for his advice on data analysis.

Correspondence concerning this article should be addressed to Jolene H. Tan, Max Planck Research Group Naturalistic Social Cognition, Max Planck Institute for Human Development, Lentzeallee 94, Berlin 14195, Germany. E-mail: tan@mpib-berlin.mpg.de

Forgiveness as an Error-Management Decision

The Theoretical Framework

Conflicts are an unavoidable aspect of social living; mechanisms that help to mitigate such conflicts are likely to be selected in animals that live in stable social groups (Aureli, Cords, & van Schaik, 2002). The forgiveness system evolved to solve these problems among humans by preserving valuable relationships after conflict and maintaining cooperative relationships over time (McCullough, Kurzban, & Tabak, 2013). Although forgiveness facilitates the victim's access to benefits that the harmdoer may be able to provide (Burnette, McCullough, Van Tongeren, & Davis, 2012), it does little to dissuade negative behaviors and may encourage further exploitation of the forgiving victim (McNulty, 2010).

We employ the term *ally* to refer to an individual with whom a relationship will bring more fitness gains than costs and employ the term *foe* to refer to the reverse. There are four possible outcomes of the forgiveness decision: Correct decisions are forgiving an ally (*true positive*) and not forgiving a foe (*true negative*), and incorrect decisions are not forgiving an ally (*false negative*) and forgiving a foe (*false positive*). To make a decision, an individual needs to estimate the *strength of evidence* that the harmdoer is an ally and set an appropriate *decision criterion* (i.e., bias) by making tradeoffs in error costs. The individual will choose to forgive if the evidence strength exceeds the decision criterion. To be efficacious in promoting decisions beneficial to fitness, the decision criterion must be sensitive to the relative cost of the two errors—it should be liberal (i.e., biased toward forgiving) when false negatives are costlier and conservative (i.e., biased toward not forgiving) when false positives are costlier (Green & Swets, 1966; Haselton & Nettle, 2006).

Relevant Information

What information is used to estimate the evidence strength and set the decision criterion? We discuss those examined in the present study.

Evidence strength. We examined four cues that may be used to estimate the harmdoer's prosocial concern for the victim's welfare: whether the harmdoer (a) had the intent to harm, (b) apologized sincerely, (c) inflicted serious harm, and (d) had committed a similar harm before (i.e., recidivism).

Three cues—intent to harm, apology, and severity—have well-established effects on forgiveness (for a review, see Fehr, Gelfand, & Nag, 2010). When the harm is intentionally committed, it suggests that the harmdoer had the goal to reduce the victim's welfare and may repeat the offense (Malle & Knobe, 1997; Petersen, Sell, Tooby, & Cosmides, 2012). A sincere apology communicates remorse and signals that the harmdoer values the relationship (Fehr & Gelfand, 2010; Ohtsubo & Yagi, 2015). The seriousness of the harm is an indication of the harmdoer's willingness to impose costs on the victim for personal benefit (Boon & Sulsky, 1997) and thus indicates the valuation of the victim's welfare (Sell, 2011).

The fourth cue—the harmdoer's recidivity—is usually investigated as a consequence of forgiveness, such as whether forgiving invites further exploitation (e.g., McNulty, 2010). Nevertheless, a harmdoer who repeatedly transgresses can be inferred to be either indifferent to or have a desire to harm the victim.

Decision criterion. Setting the criterion requires estimating whether false positives or false negatives are more costly, and individuals do so by computing the magnitude of fitness gains or losses that would result from resuming interaction with a harmdoer. The victim's cost of a false positive (erroneously forgiving a foe) can be assessed by the perceived *exploitation risk* (ER) of the harmdoer, and the cost of a false negative (erroneously not forgiving an ally) by the perceived *relationship value* (RV; Burnette et al., 2012). A liberal criterion should be adopted when the harmdoer is high in RV and low in ER (i.e., when false negatives are more costly than are false positives), whereas a conservative criterion should be adopted when it is the reverse. Our previous investigation showed that individuals indeed select their criterion according to this logic of tradeoffs (Tan et al., 2017).

Investigating the Decision Process

Decision Modeling

Decision models of the process facilitate testing of precise hypotheses about how individuals manage error costs. Predictions can be made about the presence, the direction, and also the magnitude of bias. Decision models also allow for predictions about the cognitive processes of bias as indicated by the order of information acquisition and response time. These predictions can then be tested on an individual level rather than inferred from aggregate data so as to better account for individual differences. Here, we briefly discuss two models that we used in our previous article to model forgiveness decisions—*Franklin's rule* (FR), a weighting-and-adding model, as well as *fast-and-frugal trees* (FFTs), a lexicographic heuristic—and show how they can be combined with process-tracing methodologies to investigate the decision process of forgiveness. These two models are frequently contrasted in the decision-making literature because they have different assumptions about how the mind estimates the evidence strength and implements the decision criterion, and thus the models also make different predictions about how individuals acquire information and how long a decision would take.

Franklin's rule. In FR, cue values are weighted by each cue subjective importance and then summed up to attain the evidence strength (Anderson, 1981; Gigerenzer & Goldstein, 1999). Here, the cues are binary and take the value of 1 if it is positive (i.e., strengthens the evidence strength) and 0 if it is negative (i.e., weakens). To make a decision, the evidence strength is compared to the decision criterion: forgive if the evidence strength is greater and not forgive if otherwise.

Fast-and-frugal trees. FFTs are simple decision trees that have $m + 1$ decision exits, with one exit for each of the first $m - 1$ evidence strength cues and two exits for the last cue, where m is the total number of cues (Luan, Schooler, & Gigerenzer, 2011). An exit represents a decision option (i.e., forgive or do not forgive) and is taken once a condition on the cue is met (e.g., when there was a sincere apology). The cues are ordered by importance and are considered sequentially without integration; a decision can be made after the consideration of a single cue.

The specific combination of decision exits makes up a tree's *exit structure,* which corresponds to the decision criteria (Luan et al., 2011). Unlike the decision criteria in FR, which can theoretically take any value within range, exit structure is restricted by the number of cues in an FFT and is given by the formula 2^{m-1}. Because four cues were examined in this study, there were eight possible decision criteria for FFTs (see Figure 1). To enable comparison between the models, we also applied eight matching criteria for FR (see the online supplemental materials).

Process-Tracing Tools

The use of cognitive modeling obliges researchers to make assumptions about information processing explicit and therefore facilitates predictions about both decision outcome and cognitive processes (Jarecki, Tan, & Jenny, 2015). Testing these process predictions requires supplementing data about decision outputs with those that provide a representation of the intermediate steps leading to the decision. Such process-tracing data can provide information about the cognitive steps that implement bias.

The benefits of process-tracing methods are more evident in light of the issue of model mimicry: Given a set of cue inputs, models with very different process assumptions often make the same output predictions (Willemsen & Johnson, 2011). Although mimicry alone is not a critical problem (especially when predictive accuracies are high), it makes it difficult to base conclusions about processes on prediction accuracy alone. For example, our previous study of forgiveness found that FR and FFTs not only achieved similar levels of accuracy but also made the same predictions in around 80% of the decisions (Tan et al., 2017).

To trace the process of decision-making in forgiveness, one can employ two methods commonly used in decision research: response time and pattern of information acquisition. Response time has been used as an indication of the number of pieces of information used to make a decision because it is argued that response time should increase with the number of cues used (Brandstätter, Gigerenzer, & Hertwig, 2006). Because lexicographic heuristic models such as FFTs tend to use only some of the relevant information whereas weighting-and-adding models such as FR need all information to make a decision, response time has been employed to discriminate between these two classes of models. For example, Bröder and Gaissmaier (2007) found that response times increased monotonically with the number of

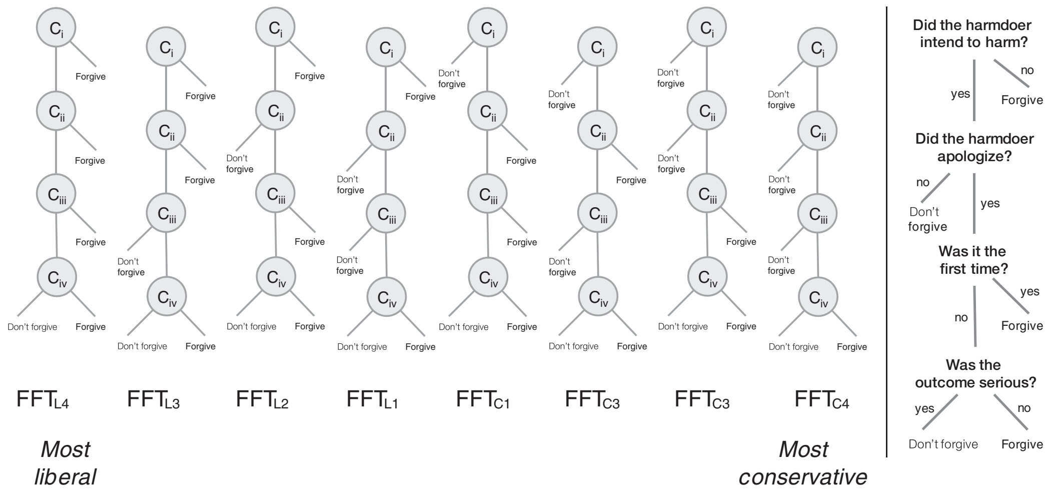

Figure 1. Fast-and-frugal trees (FFTs). The left panel displays the eight possible exit structures of an FFT, with C_i, C_{ii}, C_{iii}, and C_{iv} representing four evidence-strength cues in a fixed order. The trees are ordered from most liberal on the left to most conservative on the right. Each tree is named according to the criterion it implements: *L* indicates a liberal criterion, and *C* indicates a conservative one, and a larger number (e.g., L4) indicates a more extreme criterion than does a smaller one (e.g., L2). The right panel displays an FFT with the four evidence-strength cues investigated in our study. The exit structure of this tree implements a liberal decision criterion (i.e., FFT_{L2}). An individual using this tree will forgive if the value of the first cue is positive (i.e., there was no intent to harm), even if all other cues have negative values. If the first cue has a negative value, the individual will consider the second cue and not forgive if its value is negative (i.e., the harmdoer did not apologize).

cues that had to be searched in memory, consistent with the usage of a lexicographic heuristic. Scheibehenne, Miesler, and Todd (2007), however, did not find that response time could differentiate between a lexicographic and a weighting-and-adding strategy in a choice task where the cues were presented.

Although response time can provide a rough indication of whether more or fewer cues were used, measures of information acquisition can uncover more precisely which cues were examined and in what order. Such information acquisition patterns can provide clues about the strategy used by a decision maker, because it is assumed that information use follows soon after acquisition (Willemsen & Johnson, 2011). A common method to trace information acquisition, MouselabWEB (Schulte-Mecklenbeck, Kühlberger & Ranyard, 2011), uses mouse clicks to track the information looked up by the decision maker. In such studies, relevant information is hidden behind labeled boxes and the decision maker has to click on the box to view the information. For example, in the domain of risky choice, such techniques helped overturn the conclusions of earlier studies that relied on only response time to support the descriptiveness of models (see Brandstätter et al., 2006; E. J. Johnson, Schulte-Mecklenbeck, & Willemsen, 2008).

The Present Study

We have discussed how the forgiveness decision could be viewed as an error-management task and how tools commonly used in decision-making research can facilitate investigation of the underlying cognitive processes. In this study we extended our previous work by augmenting decision modeling with process-tracing tools that track response time and information-acquisition patterns. We looked at four evidence-strength cues instead of three in order to increase the number of potential decision criteria in FFTs and increase our ability to detect bias. We also experimentally manipulated the perception of the relative cost of errors in order to test whether individuals adjust their decision criterion accordingly.

The two decision models, FR and FFTs, were used to model each participant's hypothetical decisions in order to estimate the decision criterion adopted. We expected to replicate our previous finding that the decision criterion adopted would be influenced by the perceived cost of errors as indicated by the RV (cost of false negative) and ER (cost of false positive) of the harmdoer. As we did previously, we subtracted ER from RV to provide an index of the relative cost of errors. We expected that this measure (RV minus ER) would correlate positively with the accuracy of a model with a liberal criterion (i.e., the higher the measure, the more likely a liberal criterion of that model was adopted) and negatively with the accuracy of a model with a conservative criterion and that this effect would be strongest for models with the most extreme criteria. We also expected our manipulation to induce updates of participants' decision criteria: those whose harmdoer's ER was manipulated to be perceived as high would update their criterion to be more conservative, whereas those whose harmdoer's ER was manipulated to be low would update their criterion to be more liberal.

Even if both FR and FFTs are able to predict forgiveness decisions well, it is unclear how they would perform in predicting the decision process. Because the two models make different assumptions about cognitive implementation, they also imply different information acquisition patterns and response time. An individual using FFTs would display the following information acquisition pattern: Cues are looked up in order of importance, and cue search terminates as soon as the condition of an exit is met. On the other hand, because an individual using FR needs to weight all cues and sum them up to arrive at the evidence strength, that person would always look up all cues. In addition, an individual using an FFT would have response times that depend on the position of the decision exit taken, whereas an individual using FR should have response times that are rather consistent across trials.

Method

Participants

Two hundred ninety-eight participants (145 female; M_{age} = 34.4 years, range = 19–70) residing in the United States completed the study. Participants were recruited via Amazon's Mechanical Turk and were remunerated US$6.00. One participant admitted to not paying full attention during the study and was dropped from the analyses.

Design and Procedure

The study consisted of five phases, including a manipulation phase where participants were randomly assigned to either the high-ER or the low-ER condition. It was conducted in the MouselabWEB environment (Willemsen & Johnson, 2011), which allowed us to track information acquisition and response time. The study was approved by the ethics board of the Max Planck Institute for Human Development. More details about the method and the materials used can be found in the online supplemental materials.

Phase 1: Recall and rate. Participants were first asked to recall an incident in which they had "felt wronged, let down, betrayed, or hurt" and to spend 1–2 min writing about it. Seven participants did not write about a relevant incident and were dropped.

The recalled decision was measured with a dichotomous yes–no measure (i.e., "Did you forgive [initials of harmdoer] for what had happened?") as well as with the avoidance and revenge subscales from the Transgression-Related Interpersonal Motivations Scale (McCullough et al., 1998). The harmdoer's relationship value and exploitation risk (an indication of the relative cost of errors) were measured using the Relationship Value and Exploitation Risk (RVEX) inventory (Burnette et al., 2012) both before and after the manipulation phase. The continuous items were all measured on 7-point scale ranging from 1 (*completely disagree*) to 7 (*completely agree*).

The importance of the evidence strength cues—intent to harm, severity of offense, sincere apology, and harmdoer's recidivism—were measured using a likelihood measure. Participants rated how likely they were to forgive if the value of a cue were positive and if it were negative, independent of the other cues. They used a sliding scale ranging from 0 (*will definitely not forgive*) to 100 (*will definitely forgive*). The absolute difference between the reported likelihoods was taken as the subjective importance of that cue and used to inform the modeling procedure. Fourteen participants reported likelihoods for at least two cues in the reverse direction and were dropped because it indicates a lack of attention or understanding.

Phase 2: Hypothetical Decision Set A. Participants were asked to indicate whether they would forgive the harmdoer (yes–no) if the evidence-strength cues of the incident that they recalled had different values. Each decision trial featured a different cue profile (i.e., a permutation of evidence-strength cue values) that featured the statements used in the rating segment in Phase 1. To track information acquisition patterns, we hid cue values behind boxes labeled with the cue names, and participants had to click and hold over a box to view its value. Participants completed a total of 16 decision trials,[1] with the position of the boxes and the order of cue profiles randomized.

Phase 3: Filler task. To minimize the influence of the previous decisions on the subsequent phases, we had participants respond to 10 factual questions unrelated to forgiveness.

Phase 4: Manipulation of exploitation risk. We adapted the experimental procedure in Burnette et al. (2012) to manipulate the perception of the harmdoer's ER (i.e., cost of false positive). Participants responded to an open-ended question designed to prime the exploitative or nonexploitative aspects of the relationship with the harmdoer. Afterward, they rated the harmdoer's RV and ER again as a posttest. Three participants did not write anything meaningful and were dropped from the analyses.

Phase 5: Hypothetical Decision Set B. Participants completed another set of the trials that were identical to those in Phase 2. In both sets of trials, 19 participants did not view any cue in more than five trials and were dropped from the analyses. This left us with 254 participants (129 female; $M_{\text{age}} = 34.7$ years) included in the analyses described in the next sections.

Results

Recalled Incidents

Recalled incidents included infidelity (e.g., "She cheated on me when I was out-of-town"), lying (e.g., "I was lied to about something important"), physical assault (e.g., "He disagreed with what I was saying and decided to punch me"), and others. Participants decided to forgive in 62.2% of these incidents, and their motivation for avoidance and revenge was lower when they forgave than when they did not, Welch's $t(190.3) = 15.0$, $p < .001$, and $t(114.0) = 7.11$, $p < .001$, respectively.

Harmdoers were friends (39.4%), romantic partners (38.6%), colleagues (13.0%), family members (8.3%), and others (.8%). The average ratings of the perceived RV and ER of the harmdoer were 4.59 ($SD = 1.89$) and 3.24 ($SD = 1.41$), respectively.

Importance of Evidence-Strength Cues

The subjective importance of each cue was derived by taking the absolute difference between the likelihood rating of each cue's positive and negative statements; paired t tests indicated significant differences (all ps $< .001$ after Bonferroni correction). On average, intent was the most important ($M = 58.85$, $SD = 29.95$), followed by apology ($M = 39.47$, $SD = 22.79$), severity ($M = 35.63$, $SD = 24.44$), and recidivism ($M = 32.69$, $SD = 22.65$).

Manipulation of Relative Cost of Errors

To assess whether the manipulation procedure was successful, we examined the responses to the open-ended question as well as looked at the change in RVEX scores from the posttest to the pretest. We found that 15 participants in the low-ER condition were unable to recall any nonexploitative details and instead described exploitative details; none from the high-ER condition expressed this difficulty. Because the average change in the relative cost of errors was the same whether we excluded the 15 participants or changed their condition to high ER

[1] Due to a server error, the data for some decision trials were not recorded. Because of this, we had data for only 14 or 15 trials for 12 participants, which were analyzed in the usual way.

(M = .30, SD = .98, vs. M = .30, SD = .97), we opted for the latter. Because of the exclusion of participants described earlier, we had more participants in the high-ER condition (n = 148) than in the low-ER (n = 106).

On average, high-ER participants had a smaller increase in the relative cost of errors (RV minus ER; M = .12, SD = .93) than did low-ER participants (M = .53, SD = .98), $F(1, 252) = 11.84$, $p < .001$. However, even though the manipulation procedure was meant to alter ER, we found no difference between conditions in the change in ER ratings, $F(1, 252) = 2.41$, $p = .12$, but instead found a difference in the change in RV ratings, $F(1, 252) = 17.66$, $p < .001$. Because our hypothesis was about the change in the relative cost of errors, this was sufficient for the current purposes. Nevertheless, the crossover of effects demonstrates that despite ER and RV's being conceptually distinct, they are not orthogonal in real life; they were negatively correlated in the present study (rs = −.41 and −.45 for the pre- and posttests, respectively; both ps = .001) and also in our previous work (Tan et al., 2017). In our opinion, this merely reflects the statistical structure of the social environment and does not affect how these two variables represent the cost of errors in forgiveness.

Model Performance

We tested how well FFTs and FR would fit the decisions in the two hypothetical decision sets. In brief, we examined the accuracy of each decision criterion value for each model, analyzing the decision sets separately. For each model, we designated the criterion that was the most accurate in fitting a participant's decisions as the one that was adopted. Based on this, we calculated the overall accuracy of a model by averaging the accuracies of the designated criterion for the trials in a set. For Sets A and B, respectively, both FFTs (86.8% and 86.3%) and FR (86.3% and 85.6%) had high accuracies that were better than was the benchmark of the base rate of forgiveness[2] (55.1% and 57.3%). This replicates our previous finding (Tan, Luan, & Katsikopoulos, 2017) and supports the descriptiveness of both models for forgiveness.

Selection of Decision Criterion

We evaluated the impact of the relative cost of errors on decision criterion adopted using three separate methods. First, we found that each participant's relative cost of errors correlated positively with the fitting accuracies of models with liberal criteria and negatively with those with conservative criteria (see Figure 2); with the exception of FFT_{C1} in both sets, all ps < .01.

Second, we found that our manipulation induced an update in criterion from Decision Set A to B. In general, the accuracies of models with the two most liberal criteria increased more for the low-ER participants than for the high-ER participants, whereas the accuracies of models with conservative criteria decreased for both groups (see Figure 3). Even though the general pattern follows our predictions, the effect sizes were small and may be the result of the similarly small changes in the perceived relative cost of errors.

Third, we found that in Set B (i.e., after manipulation) the average accuracies of the models with liberal criteria were higher for low-ER than high-ER participants, whereas the accuracies of those with conservative criteria were higher for high-ER than low-ER participants (see Figure 3).

Taken together, these results conceptually replicate our previous finding (Tan et al., 2017) by showing that the decision criteria adopted by participants were affected by their perceived relative cost of errors and that participants updated their criteria when the perceived relative cost changed.

Process-Tracing Measures

On average, participants took 12.04 s (SD = 25.46) in each hypothetical decision trial. To control for individual differences in overall response latencies, we transformed these response times to z scores across trials for each decision set and for each participant. Our results support the predictions of FFTs and show that response times across trials were correlated with the number of cue lookups predicted by FFTs (r = .11, $p < .001$).

[2] The base-rate refers to the percentage of trials where the option "forgive" was chosen.

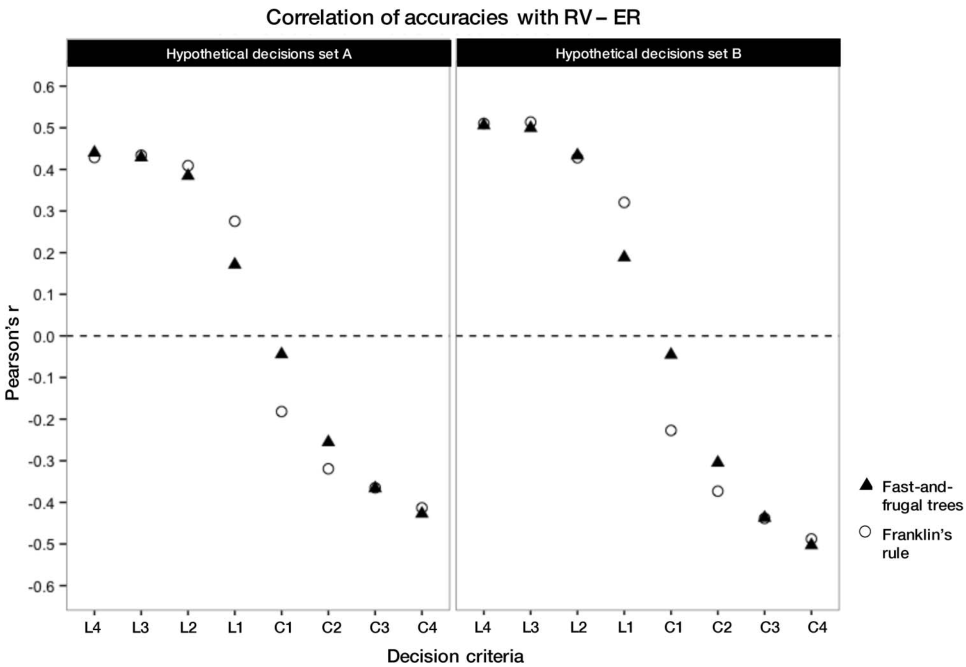

Figure 2. Relationship between the perceived relative cost of errors (i.e., relationship value [RV] minus exploitation risk [ER]) and the accuracy of a model for the two sets of hypothetical decisions. Each point in the figure shows the correlation between the perceived relative cost of errors and the accuracy of a model with a particular decision criterion. The criteria are presented from the most liberal (i.e., L4) to the most conservative (i.e., C4). With the exception of C1 of the fast-and-frugal trees for both decision sets, all $ps < .01$.

Nevertheless, response time provides only a rough indication of the number of cues used. Thus, we evaluated the information-acquisition predictions of the models by examining (a) the number of unique cues looked up, (b) which cues were looked up, and (c) the relative order of cue lookup. We also computed (d) the Levenshtein distance of the models' predictions from the data, a standard measure in computer science of the minimum number of edits (i.e., insertions, deletions, and replacements) that can transform one sequence to another (Yarkoni, Balota, & Yap, 2008). To use this measure, we treated both the predictions and data as sequences. For example, the distance between "IARS" (representing the prediction of a lookup sequence of *i*ntent, *a*pology, *r*ecidivism, and *s*everity) and "ARIS" (representing the data) is 2 because *I* had to be deleted and added before *S*. Not only can this measure account for all three aspects (i.e., number, cues, and order) but it can also account for cues being looked up more than once. We benchmarked the performance of the two models with that of a random model. More details are provided in the online supplemental materials.

Our results reported in Table 1 show that neither FFTs nor FR was better than the random benchmark in predicting which were the cues looked up or the order of cue lookup; both also had Levenshtein distances that were similar to the random benchmark. However, FR was better than both FFTs and the benchmark in predicting the number of unique cues looked up, reflecting the data showing that all four cues were looked up in 65.3% of the trials (the mean number of cues looked up was 3.41).

Even though the average number of cues looked up is high, only 20% of participants ($n = 50$) consistently looked up all four cues in all trials across both decision sets. Because this indicates that the remaining 80% of participants who did not consistently look up all four cues were unlikely to use FR, we examined the Lev-

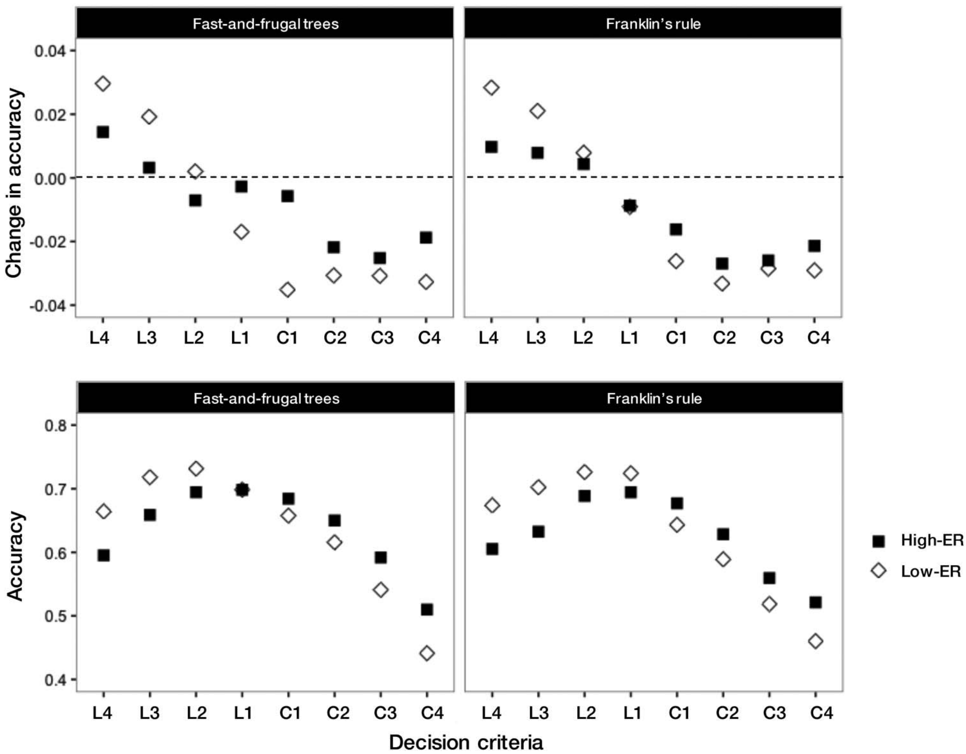

Figure 3. Results of the manipulation of exploitation risk (ER). The top panels show the average change in the accuracy of a model with a particular decision criterion (e.g., L4; note that *L* indicates a liberal criterion, and *C* indicates a conservative one, and criteria are presented from the most liberal (i.e., L4) to the most conservative (i.e., C4)) after the manipulation procedure. A positive value indicates a higher accuracy in the decision trials after the procedure (i.e., Set B) than before (i.e., Set A). The bottom panels show the average accuracies of the models in Hypothetical Decisions Set B.

enshtein distances of these two groups separately. For those 50 participants, FR had the lowest distance compared to FFTs and the random model, and this is consistent with our reasoning to examine the groups separately. Nevertheless, this was driven by the increase in the distances of FFTs and the random model (i.e., the mean distances of FFTs and the random model increased from 2.92 to 3.27 and 2.84 to 3.10, respectively) rather than a reduction in distance for FR (which changed from 2.75 to 2.86). For the remaining 204 participants, however, the main pattern of results remained—all three models had similar distances to each other and when compared with the whole sample.

Taken together, our process-tracing measures do not provide clear support for either FFTs or FR. Although the response times support the implementation assumptions of FFTs, the number of cues looked up supports FR more than FFTs.

Discussion

Understanding the forgiveness decision as an error-management task, our current investigation showed that participants' decision criteria were affected by the perceived relative cost of errors indicated by the harmdoer's perceived RV and ER, conceptually replicating our previous work and also generalizing our finding for forgiveness decisions with four, instead of three, evidence-strength cues. Our investigation was aided by the use of two decision models—FFTs and FR—that specify how the two subprocesses of EMT may be implemented cogni-

Table 1
Performance of Fast-and-Frugal Trees (FFTs) and Franklin's Rule (FR) in Predicting Information-Acquisition Patterns in Comparison With a Benchmark Random Model

	Measures			
Model	Number[a]	Cues[b]	Order[b]	Distance[c]
FFTs	1.79 (.42)	.85 (.001)	.44 (.001)	2.92 (.48)
FR	.49 (.61)	.85 (.001)	.42 (.001)	2.75 (.38)
Random	1.36 (.31)	.85 (.001)	.39 (.001)	2.84 (.42)

Note. Data presented are means, with standard errors in parentheses. FFTs = fast-and-frugal trees; FR = Franklin's rule.
[a] Absolute difference between the number of unique cues actually looked up by the participants and that of a particular model's predictions. The predictions of FFTs ranged from 1 to 4 cues, whereas the predictions of FR were always 4. [b] Values ranged from 0 to 1 and represent the proportion of correct predictions about which cues were looked up ("cues") and the relative order of the cues looked up ("order"). A value of 0 indicates that the model did not make any correct predictions, whereas a value of 1 indicates that the model's predictions were always correct. [c] Levenshtein distance is a measure of similarity between the predictions of a model about the cue lookup sequences and the actual data. It reflects the number of edits needed to transform the predictions to the data, and a higher value indicates lower similarity between the two.

tively. We found both models to have high fitting accuracies of the forgiveness decisions even though they make very different process assumptions. Moreover, we tested the models' assumptions with process-tracing tools by examining decision response times and information-acquisition patterns.

Our results do not clearly support one model over the other—response times are in favor of FFTs, whereas the number of cues looked up favor FR. However, because cues looked up may not be used in the decision but cues not looked up cannot be used in the decision, our results speak more strongly against FR. Most participants did not consistently look up all four cues, and even those who did displayed information-acquisition patterns that did not conform to the predictions of FR. It thus appears unlikely that most participants applied a weighting-and-adding strategy like FR to decide whether to forgive.

Furthermore, some methodological choices may have unwittingly increased the number of cues looked up. For instance, we opted not to make cue lookup costly as in some other studies (e.g., Molleman, van den Broek, & Egas, 2013), because costly cue search could lead to results that favor FFTs. We also followed standard guidelines to randomize the presentation order of cue boxes between trials so as to prevent order effects (i.e., Willemsen & Johnson, 2011); however, this may have created confusion and led to some boxes being opened accidentally. Follow-up work may consider imposing a small cost on cue lookup and fixing the presentation order of boxes within-subject while randomizing the presentation order between-subject.

Although we were expecting that using process measures would help us clarify which model is more descriptive of forgiveness, our results unfortunately found both models to be poor predictors of the process. Nevertheless, our predictions about cognitive processes were based on the logic of EMT and can be extended to future work that seeks to find more descriptive decision models of the forgiveness decision.

Finally, the methodology employed in this study need not be restricted to the forgiveness decision but can be extended to other error-management decisions as well (see D. D. P. Johnson, Blumstein, Fowler, & Haselton, 2013). For example, revenge decisions can be similarly understood with EMT: In deciding whether to take revenge, individuals should also consider the relative cost of errors and set the decision criterion accordingly. Thus, our approach of combining the logic of EMT with the tools from decision science can be applied to study other kinds of decisions and used in the making of high resolution maps of the adapted mind.

References

Anderson, N. H. (1981). *Foundations of information integration theory*. New York, NY: Academic Press.

Aureli, F., Cords, M., & van Schaik, C. P. (2002). Conflict resolution following aggression in gregarious animals: A predictive framework. *Animal Behaviour, 64,* 325–343. http://dx.doi.org/10.1006/anbe.2002.3071

Boon, S. D., & Sulsky, L. M. (1997). Attributions of blame and forgiveness in romantic relationships: A policy-capturing study. *Journal of Social Behavior & Personality, 12,* 19–44.

Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: Making choices without trade-offs. *Psychological Review, 113,* 409–432. http://dx.doi.org/10.1037/0033-295X.113.2.409

Bröder, A., & Gaissmaier, W. (2007). Sequential processing of cues in memory-based multiattribute decisions. *Psychonomic Bulletin & Review, 14,* 895–900. http://dx.doi.org/10.3758/BF03194118

Burnette, J. L., McCullough, M. E., Van Tongeren, D. R., & Davis, D. E. (2012). Forgiveness results from integrating information about relationship value and exploitation risk. *Personality and Social Psychology Bulletin, 38,* 345–356. http://dx.doi.org/10.1177/0146167211424582

Fehr, R., & Gelfand, M. J. (2010). When apologies work: How matching apology components to victims' self-construals facilitates forgiveness. *Organizational Behavior and Human Decision Processes, 113,* 37–50. http://dx.doi.org/10.1016/j.obhdp.2010.04.002

Fehr, R., Gelfand, M. J., & Nag, M. (2010). The road to forgiveness: A meta-analytic synthesis of its situational and dispositional correlates. *Psychological Bulletin, 136,* 894–914. http://dx.doi.org/10.1037/a0019993

Gigerenzer, G., & Goldstein, D. G. (1999). Betting on one good reason: The take the best heuristic. In G. Gigerenzer, P. M. Todd, & the ABC Research Group. (Eds.), *Simple heuristics that make us smart* (pp. 75–95). New York, NY: Oxford University Press.

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review, 109,* 75–90. http://dx.doi.org/10.1037/0033-295X.109.1.75

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York, NY: Wiley.

Haselton, M. G., Bryant, G. A., Wilke, A., Frederick, D. A., Galperin, A., Frankenhuis, W. E., & Moore, T. (2009). Adaptive rationality: An evolutionary perspective on cognitive bias. *Social Cognition, 27,* 733–763. http://dx.doi.org/10.1521/soco.2009.27.5.733

Haselton, M. G., & Nettle, D. (2006). The paranoid optimist: An integrative evolutionary model of cognitive biases. *Personality and Social Psychology Review, 10,* 47–66. http://dx.doi.org/10.1207/s15327957pspr1001_3

Jarecki, J. B., Tan, J. H., & Jenny, M. A. (2015). What is a cognitive process model? A disambiguation. *SSRN Electronic Journal*. http://dx.doi.org/10.2139/ssrn.2544831

Johnson, D. D. P., Blumstein, D. T., Fowler, J. H., & Haselton, M. G. (2013). The evolution of error: Error management, cognitive constraints, and adaptive decision-making biases. *Trends in Ecology & Evolution, 28,* 474–481. http://dx.doi.org/10.1016/j.tree.2013.05.014

Johnson, E. J., Schulte-Mecklenbeck, M., & Willemsen, M. C. (2008). Process models deserve process data: Comment on Brandstätter, Gigerenzer, and Hertwig (2006). *Psychological Review, 115,* 263–272. http://dx.doi.org/10.1037/0033-295X.115.1.263

Luan, S., Schooler, L. J., & Gigerenzer, G. (2011). A signal-detection analysis of fast-and-frugal trees. *Psychological Review, 118,* 316–338. http://dx.doi.org/10.1037/a0022684

Malle, B. F., & Knobe, J. (1997). The folk concept of intentionality. *Journal of Experimental Social Psychology, 33,* 101–121. http://dx.doi.org/10.1006/jesp.1996.1314

McCullough, M. E., Kurzban, R., & Tabak, B. A. (2013). Cognitive systems for revenge and forgiveness. *Behavioral and Brain Sciences, 36,* 1–15. http://dx.doi.org/10.1017/S0140525X11002160

McCullough, M. E., Rachal, K. C., Sandage, S. J., Worthington, E. L. J., Jr., Brown, S. W., & Hight, T. L. (1998). Interpersonal forgiving in close relationships: II. Theoretical elaboration and measurement. *Journal of Personality and Social Psychology, 75,* 1586–1603. http://dx.doi.org/10.1037/0022-3514.75.6.1586

McNulty, J. K. (2010). Forgiveness increases the likelihood of subsequent partner transgressions in marriage. *Journal of Family Psychology, 24,* 787–790. http://dx.doi.org/10.1037/a0021678

Molleman, L., van den Broek, E., & Egas, M. (2013). Personal experience and reputation interact in human decisions to help reciprocally. *Proceedings of the Royal Society: B, Biological Sciences, 280:* 20123044. http://dx.doi.org/10.1098/rspb.2012.3044

Nesse, R. M. (2005). Natural selection and the regulation of defenses: A signal detection analysis of the smoke detector principle. *Evolution and Human Behavior, 26,* 88–105. http://dx.doi.org/10.1016/j.evolhumbehav.2004.08.002

Ohtsubo, Y., & Yagi, A. (2015). Relationship value promotes costly apology-making: Testing the valuable relationships hypothesis from the perpetrator's perspective. *Evolution and Human Behavior, 36,* 232–239. http://dx.doi.org/10.1016/j.evolhumbehav.2014.11.008

Petersen, M. B., Sell, A., Tooby, J., & Cosmides, L. (2012). To punish or repair? Evolutionary psychology and lay intuitions about modern criminal justice. *Evolution and Human Behavior, 33,* 682–695. http://dx.doi.org/10.1016/j.evolhumbehav.2012.05.003

Scheibehenne, B., Miesler, L., & Todd, P. M. (2007). Fast and frugal food choices: Uncovering individual decision heuristics. *Appetite, 49,* 578–589. http://dx.doi.org/10.1016/j.appet.2007.03.224

Schulte-Mecklenbeck, M., Kühlberger, A., & Ranyard, R. (2011). The role of process data in the development and testing of process models of judgment and decision making. *Judgment and Decision Making, 6,* 733–739.

Sell, A. N. (2011). The recalibrational theory and violent anger. *Aggression and Violent Behavior, 16,* 381–389. http://dx.doi.org/10.1016/j.avb.2011.04.013

Tan, J. H., Luan, S., & Katsikopoulos, K. (2017). A signal-detection approach to modeling forgiveness decisions. *Evolution and Human Behavior, 38,* 27–38. http://dx.doi.org/10.1016/j.evolhumbehav.2016.06.004

Willemsen, M. C., & Johnson, E. J. (2011). Visiting the decision factory: Observing cognition with MouselabWEB and other information acquisition methods. In M. Schulte-Mecklenbeck, A. Kühberger, & R. Ranyard (Eds.), *A handbook of process tracing methods for decision making* (pp. 21–42). New York, NY: Psychology Press.

Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's *N:* A new measure of orthographic similarity. *Psychonomic Bulletin & Review, 15,* 971–979. http://dx.doi.org/10.3758/PBR.15.5.971

Received February 1, 2017
Revision received September 28, 2017
Accepted October 25, 2017 ■