

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/331797413>

Forgiveness Takes Place on an Attitudinal Continuum from Malevolence to Benevolence: Toward a Closer Union of Forgiveness Theory and Measurement

Preprint · March 2019

DOI: 10.31234/osf.io/wkvhn

CITATIONS

0

READS

147

9 authors, including:



Daniel E Forster
University of Miami

18 PUBLICATIONS 321 CITATIONS

SEE PROFILE



Joseph Billingsley
North Carolina State University

13 PUBLICATIONS 46 CITATIONS

SEE PROFILE



V. Michelle Russell
University of North Carolina at Greensboro

14 PUBLICATIONS 324 CITATIONS

SEE PROFILE



Michael E Mccullough
University of Miami

148 PUBLICATIONS 17,165 CITATIONS

SEE PROFILE

1
2
3
4
5
6
7
8 Forgiveness Takes Place on an Attitudinal Continuum from Malevolence to
9 Benevolence: Toward a Closer Union of Forgiveness Theory and Measurement
10
11
12
13
14
15
16
17

18 Daniel E. Forster^{1,2}, Joseph Billingsley¹, V. Michelle Russell^{3,4}, Jeni L. Burnette³, Adam
19 Smith^{5,6}, Yohsuke Ohtsubo⁵, Thomas G. McCauley¹, Joanna Schug⁷, Debra
20 Lieberman¹, & Michael E. McCullough^{1*}
21
22
23

24 ¹ University of Miami

25 ² U.S. Army Research Laboratory

26 ³ North Carolina State University

27 ⁴ University of North Carolina, Greensboro

28 ⁵ Kobe University

29 ⁶ Nagoya University

30 ⁷ College of William and Mary
31
32
33

34 *Corresponding author: Michael E. McCullough, email: mikem@miami.edu
35
36

Note: This is a non-peer-reviewed pre-print.

37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57

ABSTRACT

Researchers commonly conceptualize forgiveness as a rich complex of psychological changes involving attitudes, emotions, and behaviors. Psychometric work with the measures developed to capture this conceptual richness, however, often point to a simpler picture of the latent psychological dimensions along which forgiveness takes place. In an effort to better unite forgiveness theory and measurement, we evaluate several psychometric models for a frequently used measure of forgiveness. In doing so, we assess samples from both the United States and Japan to understand forgiveness in both non-close and close relationships. In addition, we assess the predictive utility of these models for several behavioral outcomes that traditionally have been linked to the motivations underlying forgiveness. Our results highlight models based on correlated factors models and bifactor (S-1) models. The bifactor (S-1) model evinced particular utility, identifying a single psychological dimension that spans from malevolence to benevolence while also pointing to other sources of variance that may be conceived of as method factors. The general factor of the bifactor (S-1) model consistently predicts variation in relevant criterion measures, including four different experimental economics games (when played with a transgressor), and also suffuses a second self-report measure of forgiveness. Taken together, these results suggest that forgiveness can be productively conceptualized as prosocial change along a single attitudinal continuum that ranges from malevolence to benevolence.

Keywords: forgiveness, TRIM, bifactor model, cross-cultural, reconciliation

58 INTRODUCTION

59 Because interpersonal conflict is unavoidable, forgiveness is crucial for
60 maintaining social relationships: People must be able to move past negative motivations
61 if they are to retain their valuable social partners. Indeed, socially reparative behaviors
62 have been observed in multiple primate species (De Waal & Pokorny, 2005). In
63 humans, researchers have made considerable progress in mapping the information-
64 processing mechanisms that regulate this crucial psychological process (Fehr, Gelfand,
65 & Nag, 2010). Noteworthy advances include: (i) an emerging consensus on a basic
66 definition of interpersonal forgiveness as “prosocial change toward a perceived
67 transgressor” (McCullough, Pargament, & Thoresen, 2001; p. 9); (ii) the development of
68 theoretical models of forgiveness that expand upon this basic definition by specifying a
69 conceptually rich and multidimensional set of psychological changes that can involve
70 attitudes, emotions, behaviors, and physiology (Worthington et al., 2015); and (iii) the
71 validation of multiple instruments for measuring forgiveness as conceptualized in these
72 more complex models (see Worthington et al., 2015).

73 Despite these advances, progress in understanding the psychological constructs
74 that underlie forgiveness has suffered from inconsistencies between theoretical and
75 empirical conceptualizations of forgiveness. Indeed, the research literature provides
76 multiple hints of possible misfit between *a priori* conceptions of forgiveness and how
77 measurement tools are used to model those conceptions. For example, McCullough
78 and colleagues have depicted forgiveness as a suite of motivational changes whereby a
79 victim becomes less vengeful, less avoidant, and more benevolent toward a
80 transgressor (McCullough, Fincham, & Tsang, 2003; McCullough & Hoyt, 2002). To

81 measure this three-factor conception of forgiveness, they developed the self-report
82 Transgressor-Related Interpersonal Motivations Inventory (TRIM-18; McCullough, Root,
83 & Cohen, 2006). In some of the work using this questionnaire, researchers have scored
84 the TRIM Inventory as if it reflects the operation of three distinct motivations--revenge,
85 avoidance, and benevolence (Bono, McCullough, & Root, 2008; Carmody, Gordon, &
86 Differences, 2011; Tsang, McCullough, & Fincham, 2006). Despite the supposed
87 conceptual distinctions among these three theoretical constructs, the subscales used to
88 measure them are often highly intercorrelated (e.g., McCullough, Root, & Cohen, 2006).
89 In other studies, researchers have used versions of the TRIM Inventory to model two
90 distinct motivations—the motivation to seek revenge and a bipolar motivation that
91 ranges from avoidance to benevolence (e.g., McCullough, Bono, & Root, 2007;
92 McCullough et al., 2006). In still other work, researchers have simply summed the items
93 on the TRIM Inventory as if forgiveness reflects change across a single attitudinal or
94 continuum that runs from malevolence to benevolence (Harper et al., 2014;
95 McCullough, Luna, Berry, Tabak, & Bono, 2010; McCullough, Pedersen, Tabak, &
96 Carter, 2014; Ohtsubo, Yamaura, & Yagi, 2015; Tabak, McCullough, Luna, Bono, &
97 Berry, 2012; Worthington et al., 2015). Which of these models best reflects the actual
98 psychological dimension or dimensions in which forgiveness takes place? Are there
99 really multiple distinct motivational changes underlying forgiveness, or does forgiveness
100 instead mostly reflect changes across a single underlying psychological dimension?

101 Other researchers have faced similar challenges in matching their conceptions of
102 forgiveness with the empirical realities of their measures. Subkoviak et al. (1995), for
103 example, defined forgiveness as the confluence of positive and negative cognitive,

104 affective, and behavioral components in which “a person overcomes resentment toward
105 an offender, but does not deny him/herself the moral right to such resentment. The
106 forgiver tries to have a new stance of benevolence, compassion, and even love toward
107 the offender, even though the latter has no moral right to such a response” (p. 642).
108 Subkoviak et al. sought to measure this notion of forgiveness, in all of its conceptual
109 richness, with the 60-item Enright Forgiveness Inventory, which was designed with six
110 distinct subscales to reflect both positive and negative manifestations of cognition,
111 affect, and behavior. Despite the conceptual depth of their six-dimensional model of
112 forgiveness, a single summary score based on all 60 items behaves in every way (e.g.,
113 internal consistency, test-retest reliability, correlation with other measures) similarly to
114 the six individual subscales (see also Orathinkal, Vansteenwegen, Enright, &
115 Stroobants, 2007), suggesting that the conceptual distinctions between cognitive,
116 affective, and behavioral components are not borne out empirically.

117 Similarly, Rye et al. (2001) developed their self-report measure of forgiveness
118 (“The Forgiveness Scale”) with the goal of differentiating between motivations toward
119 the offender due to the presence of positive emotions and the absence of negative
120 emotions. Despite their efforts to separate these concepts, the two subscales exhibited
121 a strong positive correlation, achieved similar internal consistency and test-retest
122 reliabilities, and evinced similar correlations with a variety of validation targets—all of
123 which suggests that the subscales may reflect a single latent continuum rather than two
124 meaningfully different dimensions of forgiveness.

125 These three sets of results suggest that the underlying attitudinal dimensions
126 upon which forgiveness takes place may be conceptually leaner than researchers have

127 heretofore posited. Indeed, they all suggest that subjects' scores on the sort of items
128 that are often used on these scales (e.g., "I'll make him or her pay"; "I'd keep as much
129 distance between us as possible"; "Even though his/her act hurt me, I still have good will
130 for him/her") are caused at least in part by a general factor that runs from negative (i.e.,
131 malevolent) evaluative reactions to the transgressor to positive evaluative reactions to
132 the transgressor, perhaps along with additional factors that influence scores on some of
133 the items (e.g., items related to revenge) but not others (e.g., items related to
134 benevolence). Since a malevolence-benevolence dimension appears to suffuse
135 interpersonal behavior in general (as exemplified in the Interpersonal Circumplex model
136 of interpersonal behavior; Gurtman, 2009), it would hardly be surprising to find that
137 forgiveness also reflected movement away from a hostile attitude regarding a
138 transgressor and movement toward a friendly or conciliatory attitude. To the extent that
139 items are caused both by such a general factor and extraneous influences (e.g.,
140 residual motives, method effects), efforts to measure forgiveness may be impeded by
141 psychometric models that cannot easily separate these two types of variance.

142 There would be additional theoretical benefit from the ability to more precisely
143 model forgiveness as a single latent continuum that runs from negative to positive
144 evaluations toward a transgressor. The prospect that the questionnaire items that
145 various researchers use to measure forgiveness (which often involve self-reports of
146 affects, cognitions, and behaviors toward a transgressor) can be neatly summarized
147 with a single general factor would suggest that forgiveness could reasonably be
148 described as a positive change in one's attitude toward an offender (inasmuch as recent
149 formal accounts depict attitudes as global evaluative reactions toward an attitude object

150 that incorporate beliefs, feelings, and behaviors; Dalege et al., 2016). This could prove
151 to be a theoretical boon for forgiveness research because so much is already known
152 about the nature of attitudes and attitude change (Albarracin & Shavitt, 2018; Bohner &
153 Dickel, 2011). Locating the concept of forgiveness within the broader conceptions of
154 attitudes and attitude change could speed future progress substantially.

155 In hopes of better uniting forgiveness theory and measurement, here we
156 systematically evaluated several modeling options for the TRIM Inventory in hopes of
157 determining the model that best depicts the underlying psychological dimension (or
158 dimensions) in which forgiveness takes place. To do so, we compared confirmatory
159 models used in the existing literature (one-factor, two-factor, and three-factor models) to
160 a model that depicts forgiveness primarily as psychological change in a single attitudinal
161 dimension (e.g., malevolence-benevolence, or hostile-friendly), along with additional
162 dimensions that might influence scores on some of the items but not others. These
163 latter models, which depict a general factor and additional item-specific content (or
164 method) factors, can be specified using variants of the bifactor modeling approach (Eid,
165 Geiser, Koch, & Heene, 2017).

166 *Bifactor Modeling: General and Specific Motives Underlying Forgiveness*

167 Recent interest in measuring general factors that span specific factors has
168 spurred the development of different so-called G-factor models (Eid et al., 2017), such
169 as bifactor models (Holzinger & Swineford, 1937). In a traditional bifactor model, a
170 general factor explains item responses across all domains while residual factors explain
171 item responses on their specified domains (Reise, 2012). However, the traditional
172 model is inappropriate when applied to typical survey methods, which has motivated the

173 development of alternative bifactor modeling approaches (Eid et al., 2017). One of the
174 models Eid et al. (2017) proposed is referred to as the bifactor (S-1), in which the
175 general factor loads on all indicators (as in the traditional bifactor model), but specific
176 (S) factors are modeled for all but one of the specific domains (hence, S-1), while the
177 non-modeled specific domain defines the scale of the general factor. When applying this
178 model to the TRIM Inventory, one might specify a model in which a malevolence-
179 benevolence factor explains item responses across domains of avoidance,
180 benevolence, and revenge, while simultaneously modeling residual factors for two of the
181 three domains, thereby leaving one domain to set the scale for the general
182 malevolence-benevolence factor.

183 To apply the bifactor (S-1) model to the TRIM Inventory, one thus faces the
184 challenge of selecting which domain best represents the general factor. We suggest
185 that avoidance is the proper reference domain because forgiveness is most consistently
186 conceptualized as a reduction in people's motives to avoid their transgressors; in other
187 words, forgiveness is conciliatory movement away from malevolence toward
188 benevolence in which victims give their transgressors a subsequent opportunity to be
189 good social partners. On this view, the motivations to re-engage (or to reduce
190 avoidance) are not necessarily benevolent or vengeful. The use of avoidance as a
191 reference domain is also supported by research indicating that approach and avoidance
192 motives are distinct, and that benevolence and revenge motives are both approach-
193 oriented, thereby making them share something important that is not shared by
194 avoidance (Elliot & Thrash, 2002). See Figure 2 for our depiction of the bifactor (S-1)
195 model as applied to the TRIM Inventory.

196 *Does the TRIM Inventory Predict Behavior toward the Transgressor?*

197 In addition to issues of specifying models that capture our understanding of
198 forgiveness as a construct, researchers who study forgiveness often lament the fact that
199 self-report measures of forgiveness are rarely validated against behavioral measures
200 that could demonstrate that they reflect interpersonally consequential psychological
201 processes (Worthington et al., 2015; for notable exceptions, see Carlisle et al., 2012;
202 Dorn, Hook, Davis, Van Tongeren, & Worthington Jr, 2014; Exline, Baumeister,
203 Bushman, Campbell, & Finkel, 2004). Without such validations against behavioral
204 measures, researchers face an uphill battle in arguing that the motivational changes
205 they measure via self-report lead to changes in how the victim actually treats the
206 offender. Indeed, on an evolutionary view of forgiveness, its function is to motivate
207 behaviors that signal to a transgressor one's willingness to re-establish cooperative
208 relations contingent on amended behavior on the transgressor's part (McCullough,
209 Kurzban, & Tabak, 2013). If the function of forgiveness is indeed to promote prosocial
210 interactions, and a self-report measure of those motivations to do so is valid, then
211 measures of that motivational change from malevolence to benevolence should
212 correlate with behavioral measures that might signal a willingness to refrain from
213 revenge and return to cooperative relations.

214 A variety of standardized experimental economics games that are generally
215 taken to reflect people's willingness to trust others, to share with others, to cooperate
216 with others in the pursuit of mutual benefit, to uphold the principle of fairness, and to
217 retaliate (or, conversely, to refrain from retaliating) have been extensively studied over
218 the past several decades. Games such as the iterated prisoner's dilemma (Axelrod,

219 1980), the Dictator Game (Forsythe, Horowitz, Savin, & Sefton, 1994), the Trust Game
220 (Berg, Dickhaut, & McCabe, 1995), and the pay-to-punish game (Fehr & Fischbacher,
221 2004) all appear to reflect the operation of a domain-general phenotype that creates
222 covariances among these games (McAuliffe, Forster, Pedersen, & McCullough, 2018;
223 Peysakhovich, Nowak, & Rand, 2014). Furthermore, the shared variance among these
224 games is reflected in peer-judgments of people's altruistic and trusting tendencies
225 (McAuliffe et al., 2018), suggesting that people's scores on these laboratory-based
226 behavioral measures reflect a friendly or prosocial approach to real-life interpersonal
227 interactions. These behavioral economic measures of prosocial behavior therefore
228 present unique opportunities to assess whether forgiveness as measured by self-report
229 is associated with affiliative interpersonal behavior toward the person whom one has
230 forgiven.

231 *The Present Project*

232 Here, we sought to assess the relative fit and predictive utility of alternative
233 psychometric models, including the bifactor (S-1) model and other more traditional
234 confirmatory factor models used for the TRIM-18 in the past. Further, we tested these
235 models using a new form of the TRIM-18 designed for non-close others (TRIM-NCO),
236 as well as the original TRIM-18. To pursue our research questions in non-close others,
237 we implemented an experimental paradigm across three settings: Online subjects from
238 the United States, laboratory subjects from the United States, and laboratory subjects
239 from Japan. In each of the three experiments, we also evaluated whether the factors we
240 estimated predicted people's scores on the iterated prisoner's dilemma game, the
241 dictator game, the trust game, and the pay-to-punish game. To then test the

242 generalizability of our modeling approaches to different relationship types and
243 measurement scales, we conducted two non-experimental survey studies across two
244 settings: Laboratory subjects from the United States and laboratory subjects from
245 Japan. In these survey studies, we used a self-report measure of forgiveness as the
246 criterion by which to assess each model's predictive utility, and we evaluated the
247 bifactor (S-1) model's ability to account for scores on other researchers' measures of
248 forgiveness as well (Hook, Worthington Jr, & Utsey, 2009; Hook, Worthington Jr, Utsey,
249 Davis, & Burnette, 2012).

250 STUDY 1 METHOD

251 *Subjects.* Subjects were 1,887 workers from Amazon's Mechanical Turk (MTurk;
252 49% female; Age: $M = 33.93$ years; $SD = 10.48$). Per our preregistration, we removed
253 all suspicious subjects from our analyses¹ ($n = 867$), resulting in a final sample of
254 1,020.² Results from analyses of all subjects (suspicious and non-suspicious) are
255 available in the supplemental materials but did not change meaningfully from those
256 reported in the main text. All subjects entered the experiment for a guaranteed \$3.50
257 with the expectation of a bonus that would depend on their decisions in the study; in
258 reality, everyone received a \$6.50 bonus and therefore earned \$10 for participating.

259 *Procedure*

¹ We presented analyses with non-suspicious subjects here to be consistent with our preregistration plan, but in the time since preregistering we learned that this practice has little empirical utility. One should analyze all data to obtain (1) more statistical power in the event that suspicious and non-suspicious subjects are statistically indistinguishable; (2) results that generalize beyond a sub-population that, for unknown reasons, would not report suspicion; and (3) results that do not violate the rules of causal inference based on experimentation (Gupta, 2011). We also note here that our results with suspicious subject included were consistent with those reported here (see supplemental materials).

²Our rates of suspicion may appear high, but we were very conservative in who we considered 'suspicious' (e.g., subjects who clearly reported suspicion because we expressly asked them about suspicion). See supplemental materials for details.

260 We used a series of experimental economic games to manipulate a
261 transgression and measure forgiveness behaviorally. Complete procedures were
262 preregistered on the Open Science Framework and are available in the supplemental
263 materials. In this manuscript, we present only the procedures that yielded the
264 information for the analyses conducted here. We conducted analyses to address
265 separate research questions in Billingsley et al. (in prep). See Table 1 for an outline of
266 the experimental design. Subjects were assigned to one of 9 conditions in a 3
267 (Message: Control, Apologetic, Aggravating) x 3 (Economic Game: Trust Game,
268 Dictator Game, Pay-to-Punish Game) between-subjects design.

269 Subjects were told that the study was designed to examine how communication
270 influences decision-making across a variety of tasks, and that they would be engaging
271 in a series of authentic interactions with another MTurk worker. To enhance the
272 believability of this interaction, we programmed the experiment using SoPHIE - the
273 Software Platform for Human Interaction Experiments (Hendriks, 2012). SoPHIE
274 enabled us to set up an online “waiting room,” which subjects entered after consenting
275 to participate. Subjects remained in the waiting room until a second subject joined,
276 whereupon the two actual subjects were paired and given the opportunity to engage in
277 an authentic communication task. Subjects who spent seven minutes in the waiting
278 room without being paired were dropped from the experiment, compensated \$0.70 for
279 their time, and permitted to participate in a successive session of the experiment. After
280 being paired, subjects were told that the purpose of the communication task was to
281 familiarize them with the program’s chat function, which would be used throughout other
282 tasks in the experiment. During the communication task, subjects took turns sending

283 and receiving five short messages without any guidance about what they should
284 discuss. Therefore, when people began engaging in experimentally manipulated
285 interactions later in the experiment (particularly during the Prisoner's Dilemma Game,
286 see below), they would have had some basis for believing the interactions were real.

287 *Iterated Prisoner's Dilemma Game*

288 Following Tabak et al. (2012), we manipulated transgressions and apologies
289 using an iterated Prisoner's Dilemma Game (PDG; called the "Decision-Making Task" in
290 this experiment). Subjects were introduced to the rules of the PDG and were told that
291 they could earn money depending on their interactions with the other subject (a pre-
292 programmed script, hereafter referred to as the "confederate"). If both subjects
293 cooperated, they both earned \$0.10 for the round; if both subjects defected, they both
294 earned \$0.05 for the round; and if one subject defected and the other cooperated, the
295 defector earned \$0.15 and the cooperator earned \$0.00. The PDG lasted 29 rounds for
296 each subject, but subjects were told that they would play for an unspecified number of
297 rounds to prevent any end-game effects. For the first 12 rounds, the confederate played
298 a generous tit-for-tat strategy: It always cooperated on the first round and, so long as
299 the subject cooperated on any round, the confederate cooperated on the subsequent
300 round. However, if the subject defected on any round, the confederate defected on the
301 subsequent round with a 50% probability.

302 Because our experimental manipulations included post-transgression apologies
303 with compensation, as a cover story we informed subjects that there would be
304 intermittent opportunities for communication throughout the decision-making task.
305 Subjects were told that they had been assigned to either a "Sender" or "Receiver" role,

306 and that the Sender could send both text and game earnings. In reality, all subjects
307 were assigned to the “Receiver” condition. As Receivers, subjects were able to respond
308 to the Sender’s messages, but were not able to return any money to the Sender.

309 *Transgressions*

310 Following the 12th round of the PDG, the game was interrupted by the first
311 opportunity for the confederate to send a message to the subject. As in Tabak et al.
312 (2012), every subject received the same message, which stated, “i think we should both
313 just hit cooperate” [sic]. Subjects were then prompted to respond. This message was
314 designed to give subjects the expectation of cooperation, regardless of the subjects’
315 strategies earlier in the game. Following the message, subjects returned to the
316 Prisoner’s Dilemma and were unconditionally defected against for 7 consecutive
317 rounds.

318 *Apology Manipulation*

319 Following the 7 rounds of unconditional defection, the confederate again sent a
320 message to the subject. For Study 1, subjects received one of three messages. In the
321 control condition, subjects read the message, “this takes more concentration than i
322 thought it would. at least it's more interesting than the HIT i did last time” [sic]. In the
323 apologetic condition, “sorry for defecting after i said cooperate. i won't do it again. i'll
324 send over some money to make it up to you” [sic]. In the aggravating condition, “sucks
325 for you, that’s just how you play the game. i’m just trying to make as much money as i
326 can” [sic]. In the apology condition, but not in the other conditions, confederates sent
327 \$1.00 of their earnings as compensation to subjects. The offer of compensation was
328 included based on evidence that apologies are most effective at influencing forgiveness

329 when the transgressor incurs some cost as part of the apology (Ohtsubo et al., 2018;
330 Ohtsubo & Watanabe, 2009; Ohtsubo et al., 2012). After the apology manipulation, the
331 confederate unconditionally cooperated for two rounds, then resumed a generous tit-for-
332 tat strategy for the remaining 8 rounds.

333 **Measures**

334 *Defections in the Prisoner's Dilemma Game*

335 We assessed forgiveness behaviorally by analyzing the ten rounds of Prisoner's
336 Dilemma Game behavior that followed the message manipulation. Rather than using
337 composite scores of the number of defections, we treated each decision in the final ten
338 rounds as an item response manifested from a latent 'propensity to defect' variable,
339 modeled as a two parameter logistic (2PL) item response model. Such a model
340 accounts for the possibility that item difficulties might actually vary across the final ten
341 decisions of the Prisoner's Dilemma. In other words, it is possible that on average
342 subjects more readily defect during earlier decisions in the sequence than during later
343 decisions. The 2PL item response model incorporates such item-based variability into
344 its estimate of each subject's propensity to defect, and should therefore yield more
345 accurate assessments than a typical composite score.

346 *One-Shot Economic Games*

347 Following the Prisoner's Dilemma Game, we randomly assigned subjects to play
348 one of three experimental economics games (Game: Trust Game, Dictator Game, Pay-
349 to-Punish Game) in a between-subjects manipulation. Each game provided an
350 opportunity to study a different forgiveness-relevant social motivation (behavioral trust,
351 behavioral benevolence, or behavioral revenge, respectively).

352 In the Trust Game (TG; Berg et al., 1995), two parties are given an endowment
353 and randomly assigned to be either 'Truster' or 'Trustee' (we used the terms 'First
354 Mover' and 'Second Mover' in our experiment). The Truster is able to send any amount
355 of the endowment to the Trustee. In turn, the Trustee receives a multiple of the amount
356 sent (three times, in our experiment). Subsequently, the Trustee is able to return any
357 proportion of the amount received (up to three times the amount sent) back to the
358 Truster. After receiving the instructions for the game and playing an example game from
359 the perspective of both roles, subjects were assigned to the role of the first mover
360 (Truster), ostensibly through random assignment.

361 In the Dictator Game (DG; Forsythe et al., 1994), the setup is very similar to the
362 Trust Game, in that two parties are given an endowment and are randomly assigned to
363 be either the 'Dictator' or 'Recipient' (we used the terms 'Decision-Maker' and
364 'Recipient' in our experiment). In the DG, the Dictator can send any amount of the
365 endowment to the Recipient at a 1:1 ratio (i.e., the amount gained by the Recipient is
366 equal to the amount paid by the Dictator). The Recipient has no opportunity to influence
367 the Dictator's decision, and the game ends once the Dictator's decision is made. After
368 receiving the instructions for the game and playing an example game from the
369 perspective of both roles, subjects were assigned to the role of the Decision-Maker
370 (Dictator), ostensibly through random assignment.

371 In the Pay-to-Punish Game (PTPG; Fehr & Fischbacher, 2004), the setup is
372 nearly identical to the DG, in that two parties are given an endowment and randomly
373 assigned to be either 'Decision-Maker' or 'Recipient'. In the PTPG, the Decision-Maker
374 is able to spend any amount of the endowment to remove some amount from the

375 Recipient's endowment (in our experiment, the amount removed was four times the
376 amount paid). As in the DG, the Recipient is not given an opportunity to influence the
377 Decision-Maker, and the Decision-Maker cannot make any financial gains by removing
378 money from the recipient. After receiving the instructions for the game and playing an
379 example game from the perspective of both roles, subjects were assigned to the role of
380 the Decision-Maker, ostensibly through random assignment.

381 *Self-Report Measure of Forgiveness: Transgression Related Interpersonal Motivations*
382 *Inventory for Non-Close Others*

383 Immediately following subjects' decisions in the second experimental economic
384 game, subjects completed the 18-item Transgression Related Interpersonal Motivations
385 Inventory for Non-Close Others (TRIM-NCO; see Appendix A). Typically, researchers
386 have measured forgiveness using items that assume an existing relationship between
387 victim and transgressor, thereby limiting our ability to assess forgiveness between
388 people who are interacting for the first time (as in a typical laboratory context). The
389 TRIM-NCO Inventory is designed to overcome this limitation by re-wording items from
390 the original TRIM-18 Inventory so that they are sensible to people who are rating their
391 attitudes toward strangers.

392 Although we use the TRIM-NCO Inventory as a predictor of behavioral indicators
393 of willingness to restore cooperative relations, we avoided assessing self-report
394 forgiveness prior to our behavioral measures to prevent any contamination effects. We
395 based the items in the TRIM-NCO on a prior measure, the TRIM-18, which McCullough
396 et al. (2006) developed for measuring interpersonal forgiving in close relationships. As
397 with the traditional TRIM-18, the TRIM-NCO provides a self-report measure of the

421 revenge), and a single factor ('forgiveness')³. We also analyzed the TRIM-NCO using a
422 bifactor (S-1) model, in which a general factor explains variation in all item responses
423 while specific factors explain unique variation among subsets of items.

424 Model fit for the alternative factor structures are available in Table 2a. The three-
425 factor and bifactor (S-1) models fit best, with the fit of these two models being
426 indistinguishable under reasonable constraints (Geiser, Eid, & Nussbeck, 2008);
427 therefore, we opted to compare models of forgiveness as bifactor (S-1) and three factor
428 models. Path coefficients for each of these models are displayed in Table 3a.

429 We also concluded that the item, "I hope he/she gets what he/she deserves,"
430 which was designed as a revenge indicator, performed poorly at both the general and
431 specific levels. This was demonstrated by its abnormally low factor loadings across all
432 of the models we tested. Our interpretation of this result is that the item is inherently
433 ambiguous—people believe that those who behave prosocially deserve good things and
434 should get what they deserve. In other words, the negative connotations typically
435 associated with 'just deserts' is not necessarily clear in the way the item was phrased.
436 Therefore, this item was excluded from the analyses presented here but we also report
437 results with this item included in the supplemental materials.

438 *Did a Three-Factor Model Predict Economic Game Behavior?*

439 We regressed subjects' propensities to defect and their one-shot economic game
440 behavior (TG, DG, or PTPG) on each of the three forgiveness factors (avoidance,
441 revenge, and benevolence). We report standardized coefficients here and in Table 3a.

³ Although previous one-factor models used a Rasch model for graded responses, we used a less restrictive graded response model (Samejima, 1969) for comparison, which does not constrain all item loadings and threshold parameters to be equal to each other.

442 Propensity to defect in the final ten rounds of the PDG was predicted by
443 avoidance, $b = -.409$, $se = .097$, $p < .001$, but not by revenge, $b = -.109$, $se = .073$, $p =$
444 $.134$, or by benevolence, $b = -.011$, $se = .131$, $p = .932$. Amount transferred in the TG
445 was predicted by benevolence, $b = .479$, $se = .181$, $p = .008$, but not by avoidance, $b = -$
446 $.048$, $se = .130$, $p = .712$, or by revenge, $b = -.121$, $se = .107$, $p = .258$. Amount sent in
447 the DG was not predicted by any of the three factors (avoidance: $b = .240$, $se = .186$, $p =$
448 $.198$; revenge: $b = .157$, $se = .124$, $p = .207$; benevolence: $b = .004$, $se = .241$, $p =$
449 $.985$). Finally, the amount removed in the PTPG was predicted by revenge, $b = -.462$, $se =$
450 $.096$, $p < .001$, but not by avoidance, $b = -.074$, $se = .141$, $p = .599$, or by
451 benevolence, $b = .157$, $se = .176$, $p = .373$.

452 Overall, one could argue that the three-factor model excels at determining which
453 motivations are underlying different types of experimental-economic behavior. However,
454 three distinct factors are unable to capture whether a common process is underlying
455 some aspect of these behaviors. Moreover, the three-factor model leads to ambiguity in
456 interpreting why none of the factors predicted transfers in the DG: this result could
457 indicate either that *none* of the three motives are related to transfers in the DG or that *all*
458 of the motives relate to DG transfers but explain overlapping variance. Therefore, we
459 compared these results to those from a bifactor (S-1) model, which we used to extract a
460 general factor (forgiveness, scaled by avoidance) as well as specific factors for
461 benevolence and revenge.

462 *Did a Bifactor (S-1) Model Predict Economic Game Behavior?*

463 We specified the bifactor (S-1) model with the subset of avoidance items as the
464 reference domain. The general factor was scaled using the item 'I would not trust

465 him/her' (reverse-scored), the specific domain for benevolence was scaled using the
466 item 'I would have good will for him/her', and the specific domain for revenge was
467 identified with 'I would want to seek revenge' (reverse-scored).

468 We then regressed subjects' propensities to defect and their one-shot economic
469 game behavior (TG, DG, or PTPG, depending on their randomized condition) on the
470 general forgiveness factor (as scaled by reverse-scored 'avoidance', making positive
471 values represent greater forgiveness) and two specific factors (reverse-scored revenge
472 and benevolence) in the bifactor (S-1) model. We report standardized coefficients here
473 and in Table 3a. Propensity to defect in the PDG was predicted by the general factor, b
474 $= -.493$, $se = .035$, $p < .001$, but not by the specific factors representing revenge
475 (reverse-scored), $b = -.080$, $se = .056$, $p = .152$, or benevolence, $b = -.011$, $se = .058$, p
476 $= .853$. Amount transferred in the TG was predicted by the general factor, $b = 0.293$, se
477 $= .051$, $p < .001$, and by the benevolence specific factor, $b = .253$, $se = .084$, $p = .003$,
478 but not by the revenge specific factor, $b = -.110$, $se = .083$, $p = .185$. Amount sent in the
479 DG was predicted by the general factor, $b = .350$, $se = .049$, $p < .001$, but not by either
480 the revenge specific factor, $b = .161$, $se = .096$, $p = .093$, or the benevolence specific
481 factor, $b = -.061$, $se = .085$, $p = .475$. Finally, the amount paid to punish in the PTPG
482 was predicted by the general factor, $b = -.218$, $se = .056$, $p < .001$, and by the revenge
483 specific factor, $b = -.343$, $se = .074$, $p < .001$, but not by the benevolence specific factor,
484 $b = .027$, $se = .078$, $p = .729$.

485 Here, we see that the general factor from the bifactor (S-1) model predicted
486 meaningful variation in all of our behavioral outcomes, whereas the factors from the
487 correlated model performed less consistently. Also, none of the factors from the

488 correlated model were able to explain meaningful variation in Dictator Game transfers;
489 not because the factors were unrelated, but because the correlated factors explained
490 overlapping variation in that outcome, which was made apparent by the general factor's
491 success in predicting DG transfers in the bifactor (S-1) model.

492 Finally, we correlated the general factor with a score derived by simply taking the
493 mean of all 17 items from the TRIM-NCO. The two measures were highly correlated, $r =$
494 $.946$, $p < .001$. Unsurprisingly, this simple 17-item composite was also significantly
495 correlated with scores on each of the four experimental economics games (see Table
496 5a). Jointly, these results suggest that the general factor can be estimated with
497 reasonable fidelity simply by taking the mean of all 17 items on the TRIM-NCO.

498 *Did Item Responses Support a Continuum from Malevolence to Benevolence on the*
499 *General Factor?*

500 To better understand how the general factor corresponds to our
501 conceptualization of forgiveness, we computed location indices (LIs) for each item
502 based on the item response function (LI_{IRF} ; Ali, Chang, & Anderson, 2015). The LI_{IRF}
503 value represents the difficulty of a polytomous item (i.e., the probability of endorsing a
504 higher value on the scale) for a person with a specified factor score on the latent
505 construct (we based LI_{IRF} values on a latent score of 0, but the pattern of findings would
506 hold across any chosen latent score). Although one could summarize each item's
507 difficulty by computing the mean or median threshold for each item, the LI_{IRF} tends to
508 characterize item difficulty better than a crude central tendency measure because it
509 integrates information from the item's loading and all of its thresholds. We present LI_{IRF}
510 values in Table 5a, ordered from lowest (i.e., 'easiest') to highest (i.e., 'hardest'). As

511 shown in the table, the LI_{IRF} values are arrayed along the general factor with revenge
512 items on the easier end and a mixture of avoidance and benevolence items on the more
513 difficult end, suggesting that a motivational or attitudinal continuum that stretches from
514 malevolence to benevolence underlies the general factor. We interpret this descriptive
515 pattern to suggest that people with very low scores on the general factor can renounce
516 revenge whereas they require higher scores on the general factor to endorse approach
517 and benevolence.

518 STUDY 1 DISCUSSION

519 For Study 1, we collected a large online sample to test the relative fit and
520 predictive utility of alternative TRIM-NCO models. Based on our results, we found that
521 the three-factor and bifactor models of the TRIM-NCO fit best. Therefore, we used the
522 results of these two measurement models to predict outcomes from four experimental
523 economics games that measure forgiveness-relevant social motivations (Billingsley &
524 Losin, 2017): an iterated prisoner's dilemma game and three one-shot economic games
525 (trust game, dictator game, and pay-to-punish game). The three factors that are
526 traditionally used to account for the covariances among the items on the TRIM
527 (revenge, avoidance, and benevolence) were each significantly associated with scores
528 on one of the four experimental economic games we examined. However, the general
529 factor we derived from the bifactor (S-1) model of those covariances was significantly
530 associated with scores on all four economic games, and in the directions one would
531 expect if that general factor did indeed reflect a latent attitudinal continuum that runs
532 from antagonism to friendliness, or malevolence to benevolence). The group factors
533 representing variance specifically attributable to the revenge items and the benevolence

534 items predicted significant amounts of unique variance in two of the experimental
535 economics games we examined, which suggests that they may possess incremental
536 validity as measures of forgiveness. However, the consistent usefulness of the general
537 factor for predicting all four behavioral outcomes suggests that it may be the factor that
538 most reliably uncovers the behaviorally relevant motivational changes underlying
539 forgiveness. We also found that this general factor can be dependably estimated simply
540 by taking the mean of all 17 of the items on the TRIM-NCO.

541 STUDY 2

542 Researchers' inability to control data collection conditions in online samples has
543 caused much concern among researchers (Buchanan & Scofield, 2018; Buhrmester,
544 Kwang, & Gosling, 2011; Kan & Drummey, 2018). For example, scholars have noticed
545 that substantial numbers of online research subjects provide false responses or find
546 ways of participating multiple times, thereby clouding our understanding of the effects
547 the wish to study (Buchanan & Scofield, 2018; Kan & Drummey, 2018). To ensure that
548 Study 1's results were generalizable beyond online samples, we conducted Study 2
549 using a typical laboratory setting at a university in the United States.

550 STUDY 2 METHOD

551 *Subjects*

552 To maximize statistical power and access to subjects in Study 2, we recruited
553 subjects from two sources. First, we recruited 342 students from the undergraduate
554 psychology subject pool at a large university in the southeastern United States (Sex not
555 collected; Age: $M = 19.41$ years; $SD = 4.77$). Although subjects initially believed they
556 could earn up to \$10 (implying the possibility of earning less), all subjects received the

557 full \$10 in addition to partial course credit for participating. Second, we recruited 79
558 subjects from the surrounding community via email and Craigslist. Subjects from the
559 community sample were on average 36.20 ($SD = 11.95$) years old. All community
560 subjects received a \$20 show-up fee. In addition, these subjects initially believed they
561 would earn a bonus up to \$11 (depending on their decisions in the study), although all
562 community subjects received the full \$11, resulting in a total of \$31 in compensation for
563 their participation. As in Study 1, we excluded subjects who indicated some level of
564 suspicion ($N = 95$). We also excluded responses from one of the conditions (see
565 procedural differences below for details), for a final sample of 228, with analyses of all
566 subjects available in the supplemental materials.

567 *Procedure*

568 For Study 2, the student and community samples completed procedures similar
569 to those of Study 1. Upon arriving at the laboratory, subjects were seated at a computer.
570 After subjects provided informed consent, a researcher instructed them that they would
571 be playing economic games with an anonymous partner located elsewhere on campus
572 via a computer network. The remainder of the experiment was conducted via computer
573 using the E-Prime 2.0 software (Psychology Software Tools, 2012).

574 *Procedural Differences Between Study 2 and Study 1*

575 In addition to its laboratory setting and its implementation in E-Prime (vs.
576 SoPHIE), Study 2 diverged from Study 1 by including a No-Transgression control
577 condition ($N = 98$). Cases from the No-Transgression control condition were excluded
578 from analyses presented here because we were interested in understanding the utility of

579 the TRIM-NCO when a transgression occurred. Analyses including the no-transgression
580 condition appear in the companion to this manuscript (Billingsley et al., in prep).

581 Study 2 also diverged from Study 1 in that we had subjects engage in only one
582 the economic games we examined in Study 1 (the Trust Game). We chose the Trust
583 Game for its face-valid relevance to a continued interaction, whereas the other two
584 games were one-sided interactions.

585 STUDY 2 RESULTS

586 We followed the same analysis procedures as we did in Study 1. Model fit
587 statistics for the alternative factor structures are available in Table 2a. As in Study 1, the
588 three-factor and bifactor (S-1) models fit best, so we proceeded to evaluate the
589 predictive utility of the two modeling approaches. Model specification procedures were
590 identical to those of Study 1. Also consistent with Study 1, we removed the item, 'I hope
591 he/she gets what he/she deserves,' from the analyses reported here.

592 *Did a Three-Factor Model Predict Economic Game Behavior?*

593 We regressed subjects' propensities to defect and their Trust Game behavior on
594 each of the three forgiveness factors (avoidance, revenge, and benevolence). We report
595 standardized coefficients here, as well as in Table 3a.

596 Propensity to defect in the prisoner's dilemma game was not predicted by any of
597 the three factors (avoidance: $b = -.250$, $se = .133$, $p = .060$; revenge: $b = -.154$, $se =$
598 $.122$, $p = .209$; benevolence: $b = -.067$, $se = .134$, $p = .617$). Amount transferred in the
599 trust game was predicted by (reverse-scored) revenge, $b = -.287$, $se = .107$, $p = .007$,
600 and benevolence, $b = .358$, $se = .117$, $p = .002$, but not by avoidance, $b = .187$, $se =$

601 .127, $p = .140$. These results are in contrast to Study 1, in which defections in the PDG
602 were predicted by avoidance and transfers in the TG were predicted by benevolence.

603 *Did a Bifactor (S-1) Model Predict Economic Game Behavior?*

604 We then regressed subjects' propensities to defect and their Trust Game
605 behavior on the general forgiveness factor and the two specific factors (revenge and
606 benevolence) in the bifactor (S-1) model.

607 Propensity to defect in the PDG was predicted by the general factor, $b = -.403$, se
608 $= .062$, $p < .001$, but not by the specific factors representing revenge, $b = -.101$, $se =$
609 $.092$, $p = .274$, or benevolence, $b = -.056$, $se = .094$, $p = .554$. Amount transferred in the
610 TG was predicted by the general factor, $b = .263$, $se = .064$, $p < .001$, by the revenge
611 specific factor (reverse-scored), $b = -.235$, $se = .078$, $p = .002$, and by the benevolence
612 specific factor, $b = .224$, $se = .081$, $p = .005$. As in Study 1, these results suggest that
613 the general factor from the bifactor (S-1) model captures behaviorally relevant variance
614 in people's regard for someone who has recently harmed them.

615 Finally, we correlated the general factor with a score derived from simply taking
616 the mean of all 17 items from the TRIM-NCO. As in Study 1, the two measures were
617 highly correlated, $r = .937$, $p < .001$. Unsurprisingly, this simple 17-item composite was
618 also significantly correlated with scores in the Trust Game (see Table 4a). Jointly, these
619 results suggest that the general factor can be estimated with reasonable fidelity simply
620 by taking the mean of all 17 items on the TRIM-NCO.

621 *Did Item Responses Support a Continuum from Malevolence to Benevolence on the*
622 *General Factor?*

623 As in Study 1, we computed item difficulty values using Ali et al.'s (2015) LI_{IRF}
 624 method, which are presented in Table 5a. Again, we found that LI_{IRF} values were sorted
 625 along the general factor with revenge items on the easier end and a mixture of
 626 avoidance and benevolence items on the more difficult end. Further, LI_{IRF} values for
 627 each item were strongly correlated with those found in Study 1, $r = .85$, as were the
 628 rankings themselves, $r = .61$, indicating consistency across samples and providing
 629 further support that the general factor represents an attitudinal continuum that ranges
 630 from malevolence (revenge motivation) to benevolence.

631 STUDY 2 DISCUSSION

632 Here, we largely replicated the results found in Study 1, in that the general factor
 633 and the group factors from the bifactor model consistently predicted subjects' behavior
 634 in the prisoner's dilemma and the Trust Game. We note that the three-factor correlated
 635 model yielded little consistency across Studies 1 and 2. This is in contrast to the pattern
 636 observed using the bifactor (S-1) modeling approach: The general factor predicted
 637 scores on both of the games studied, and the revenge and benevolence factors both
 638 predicted unique variance in subjects' TG transfers (but not in their PDG scores). Oddly,
 639 however, the correlation between the (reverse-scored) revenge-specific factor and trust
 640 game scores was negative rather than positive—subjects who were more vengeful than
 641 their scores on the general factor (scaled by avoidance) would indicate sent more
 642 money in the trust game. It is tempting to attribute the negative relationship between two
 643 prosocial constructs to the vagaries of sampling error variance and to conclude that it
 644 casts suspicion upon the validity of the group factors as measures of forgiveness.

645 We also found that the general factor can be dependably estimated simply by
646 taking the mean of all 17 of the items on the TRIM-NCO.

647 STUDY 3

648 Psychology has faced pointed criticisms for over-reliance upon samples of
649 Western undergraduates. The material wealth, extensive education, democratic values,
650 and highly industrialized backgrounds of these students relative to much of the world
651 call into question the representativeness of results obtained using such samples
652 (Henrich, Heine, & Norenzayan, 2010). Although Study 1 draws from an adult rather
653 than an undergraduate sample, and Study 2 includes community members as well as
654 students, our results thus far demonstrate the utility of the bifactor (S-1) model for the
655 TRIM-NCO measure in a Western—specifically, American—context. To increase the
656 generalizability of our results, we attempted to replicate our findings in Japan, a culture
657 known to differ significantly from that of the United States along several dimensions
658 pertinent to forgiveness, notably relational mobility and collectivism vs. individualism
659 (Kashima et al., 1995; Schug, Yuki, Horikawa, & Takemura, 2009; Yamagishi &
660 Yamagishi, 1994). If successful, our efforts would furnish researchers with preliminary
661 evidence that the usefulness of the bifactor (S-1) for analyzing TRIM data extends
662 beyond the geographical borders of the United States and beyond the linguistic
663 boundaries of the English language. Such efforts would provide researchers from
664 diverse cultures—perhaps especially those cultures where a translated version of the
665 TRIM-18 is already in use—with a basis for modeling forgiveness as prosocial change
666 that takes place on a latent attitudinal continuum that runs from hostility to friendliness.

667 STUDY 3 METHOD

668 *Subjects*

669 300 Japanese citizens were recruited using a Japanese crowdsourcing service,
670 Lancers, Inc (65% female; Age: $M = 36.57$; $SD = 10.12$). As in the previous
671 experiments, we excluded subjects who indicated any level of suspicion ($N = 114$)
672 before conducting analyses, resulting in a final sample of $N = 186$. Results with all
673 subjects included are available in the supplementary materials.

674 *Procedures*

675 As in Study 1, Study 3 was programmed using SoPHIE (Hendriks, 2012). All
676 procedures were identical between Studies 1 and 3, except that subjects played only
677 the Trust Game following the Prisoner's Dilemma Game, much like we did in Study 2,
678 with endowments identical to those in Study 1. Study 3 also differed from Study 1 in
679 how the chat was implemented. Specifically, instead of sending and receiving five open-
680 ended chat messages, subjects simply typed the Japanese word for 'hello' and received
681 a commensurate 'hello' from the other subjects. We made this change because
682 Japanese subjects had difficulty chatting with a stranger under the constraints of the
683 chat function, which prevented them from moving forward with the study.

684 *Translation*

685 A Japanese version of the TRIM-18 translation already existed (Ohtsubo,
686 Yamaura, & Yagi, 2015), so we modified this existing translation to make the items
687 more applicable to non-close others, just as we did with the existing English version for
688 Studies 1 and 2.

689

STUDY 3 RESULTS

690 We followed the same analysis procedures as we did in Studies 1 and 2. Model
691 fit statistics for the alternative factor structures are available in Table 2a. As in Study 1,
692 the three-factor and bifactor (S-1) models fit best, so we proceeded to evaluate the
693 predictive utility of the two modeling approaches. Model specification procedures were
694 identical to those of Studies 1 and 2. Also consistent with Studies 1 and 2, we removed
695 the item, 'I hope he/she gets what he/she deserves,' from the analyses reported here.

696 *Did a Three-Factor Model Predict Economic Game Behavior?*

697 We regressed subjects' propensities to defect and their one-shot economic game
698 behavior in the trust game on each of the three forgiveness factors (avoidance,
699 revenge, and benevolence). We report standardized coefficients here and in Table 3a.

700 Propensity to defect in the PDG was predicted by benevolence, $b = -.441$, $se =$
701 $.178$, $p = .013$, but not by avoidance, $b = -.099$, $se = .178$, $p = .580$, or revenge, $b =$
702 $.178$, $se = .155$, $p = .250$. Amounts transferred in the TG were not predicted by any of
703 the three factors (avoidance: $b = -.030$, $se = .158$, $p = .850$; revenge: $b = .102$, $se =$
704 $.128$, $p = .426$; benevolence: $b = -.007$, $se = .159$, $p = .965$).

705 *Did a Bifactor (S-1) Model Predict Economic Game Behavior?*

706 We then regressed subjects' propensity to defect in the prisoner's dilemma game
707 and their scores on the trust game on the general forgiveness factor and two specific
708 factors (revenge and benevolence) in the bifactor (S-1) model.

709 Propensity to defect in the PDG was significantly associated with the general
710 factor, $b = -.379$, $se = .066$, $p < .001$, but not with the specific factors representing
711 revenge (reverse-scored), $b = .139$, $se = .120$, $p = .244$, or benevolence, $b = -.148$, $se =$
712 $.097$, $p = .126$. As was the case with the standard three-factor model for the TRIM,

713 amounts transferred in the trust game were not significantly associated with the general
714 factor ($b = .030$, $se = .070$, $p = .669$), the revenge specific factor ($b = .077$, $se = .100$, p
715 $= .442$) or the benevolence specific factor: $b = -.002$, $se = .091$, $p = .986$). Although we
716 found that the general factor was able to predict subjects' propensities to defect in the
717 PDG, we were unable to explain any of the variation in TG transfers with any of the
718 factors from either the correlated or bifactor (S-1) models.

719 Finally, we correlated the general factor with a score derived from simply taking
720 the mean of all 17 items from the TRIM-NCO. The two measures were highly correlated,
721 $r = .940$, $p < .001$. Unsurprisingly, the simple 17-item composite was also significantly
722 correlated with scores on the prisoner's dilemma game, but not with scores on the trust
723 game (see Table 4a). Jointly, these results suggest that the general factor can be
724 estimated with high fidelity with the mean of all 17 items on the TRIM-NCO.

725 *Did Item Responses Support a Continuum from Malevolence to Benevolence on the*
726 *General Factor?*

727 As in Studies 1 and 2, we computed LI_{IRF} for each item (Ali et al., 2015), which
728 are presented in Table 5a. Again, we found that LI_{IRF} values were sorted along the
729 general factor with revenge items on the easier end and a mixture of avoidance and
730 benevolence items on the more difficult end. We also found that LI_{IRF} values and LI_{IRF}
731 ranks for each item were strongly correlated with those found in Study 1 ($r = .923$ and $r =$
732 $.838$, respectively) and those found in Study 2 ($r = .757$ and $r = .600$, respectively),
733 providing further evidence for consistency across samples and for the interpretation that
734 the general factor represents a malevolence-benevolence continuum.

735

STUDY 3 DISCUSSION

736 In Study 3, we obtained similar results regarding the factor structure of the TRIM-
737 NCO Inventory, with the bifactor model and three factor models exhibiting good model
738 fit. As was the case in Studies 1 and 2, the general factor from the bifactor (S – 1)
739 model was associated with scores on the prisoner’s dilemma game. So, too, was one of
740 the subscales that resulted from a traditional scoring of the TRIM. None of the factors
741 derived from the standard three-factor model or the bifactor (S – 1) model predicted
742 scores in the Trust Game. Because the general factor was significantly associated with
743 scores on every other experimental economics game from all three experiments, the
744 non-significant association of forgiveness with scores on the trust game may reflect a
745 true cultural difference rather than the result of sampling error. Finally, a mean of all 17
746 TRIM items performed very much like the general factor, suggesting that a simple
747 composite of all 17 items is a reasonable way to estimate the general factor.

748 STUDY 4

749 In Studies 1-3, we probed the hypothesis that a single dimension spanning
750 malevolence and benevolence underlies forgiveness. We used the TRIM-NCO, a new
751 version of widely-used TRIM-18 (McCullough, Cohen, & Root, 2006) that was modified
752 to assess interpersonal motivations between individuals encountering one another for
753 the first time. However, forgiveness often occurs in close interpersonal relationships
754 rather than in first-time, anonymous encounters, with great significance for our personal
755 lives and well-being (Rusbult, Hannon, Stocker, & Finkel, 2005).

756 To expand our findings beyond first-time interactions between strangers, we
757 therefore report the results of a fourth study, in which we applied the bifactor (S-1)
758 modelling technique to the more traditional 18-item Transgression Related Interpersonal

759 Motivations Inventory (TRIM-18; McCullough, Cohen, & Root, 2006), which was
760 designed to measure forgiveness following a transgression committed by a familiar
761 person in a real-world setting. In addition to expanding on Studies 1-3 by assessing
762 forgiveness in close relationships, we also sought to generalize our findings to self-
763 report instruments beyond variants of the TRIM. To do so, we used subjects' responses
764 from the Decision to Forgive and Emotional Forgiveness Scales (Hook et al., 2009;
765 Hook et al., 2012). These additional scales allowed us to determine whether a general
766 factor reflects a general propensity to forgive that suffuses the items from other
767 measures as well.

768 STUDY 4 METHOD

769 *Subjects*

770 Subjects were 168 undergraduate students recruited from introductory
771 psychology classes at a Southeastern University in the United States (distinct from that
772 of Study 2; 59.5% female; Age not collected). Subjects who were of East Asian
773 nationality were excluded so as not to confound a broader cross-national project design
774 (see Study 5). Subjects completed the survey outside of the lab in exchange for course
775 credit. Unlike Studies 1-3, study 4 utilized a correlational rather than experimental
776 design. As a result, our experiment did not involve deception, we did not probe subjects
777 for suspicion, and no subjects were excluded from analyses.

778 *Procedure*

779 All data were collected using an online Qualtrics survey with standardized
780 instructions embedded in the survey. To study forgiveness in real-world transgressions,
781 we instructed subjects to “[t]hink of a time that a close other person did something to

782 upset you, hurt you, or otherwise commit an offense that caused a rift in your
783 relationship.” To increase the salience of the memory, we asked subjects to describe
784 the context and outcome of the offense. Immediately after describing the transgression,
785 subjects completed several self-report measures, including the TRIM-18, Decision to
786 Forgive Scale, and Emotional Forgiveness Scale. Subjects also completed other
787 measures beyond the scope of this paper.

788 **Measures**

789 *Decisional Forgiveness*

790 The Decision to Forgive Scale (DFS) measures the degree to which the victim of
791 a transgression deliberately works to replace negative behavior towards a transgressor
792 with positive, prosocial behavior (Hook et al., 2009; Hook et al., 2012). Subjects were
793 instructed as follows: “Think of your current intentions toward the person who hurt you.
794 Indicate the degree to which you agree or disagree with the following statements.” For
795 example, subjects were asked how much they agree with the statement, “I will not talk
796 with him or her.” The DFS is an 8-item scale with response options ranging from 1
797 (“Strongly Disagree”) to 5 (“Strongly Agree”; see Appendix C). We scored subjects’ DFS
798 responses so that higher scale scores indicated greater decisional forgiveness.

799 *Emotional Forgiveness*

800 The Emotional Forgiveness Scale (EFS) measures the degree to which a victim
801 replaces negative emotions towards a transgressor (e.g., anger) with positive emotions
802 (e.g., compassion; Hook et al., 2009). Subjects were instructed as follows: “Think of
803 your current emotions toward the person who hurt you. Indicate the degree to which you
804 agree or disagree with the following statements.” For example, subjects were asked

805 how much they agree with the statement, “I no longer feel upset when I think of
806 him/her.” The EFS is an 8-item scale with response options ranging from 1 (“Strongly
807 Disagree”) to 5 (“Strongly Agree”; see Appendix D). We scored the EFS so that higher
808 scale scores indicated greater emotional forgiveness.

809 *Transgression-Related Interpersonal Motivation Scale (Close Others)*

810 Forgiveness motivations were measured using the 18-item Transgression-
811 Related Interpersonal Motivation Scale for close others (TRIM-18; McCullough, Root, &
812 Cohen, 2006). Subjects were instructed as follows: “For the following questions, please
813 indicate your current thoughts and feelings about the person who hurt you; that is, we
814 want to know how you feel about that person right now. Next to each item, circle the
815 number that best describes your current thoughts and feelings.” Responses ranged
816 from 1 (“Strongly Disagree”) to 5 (“Strongly Agree”). Consistent with Studies 1-3,
817 negatively worded items (i.e., items indicating less forgiveness) were reverse scored so
818 that higher scores indicated greater forgiveness. For example, strong agreement with
819 the items “I will make him or her pay” from the revenge scale and “I withdraw from
820 him/her” from the avoidance scale contributed to lower scores on the TRIM-18. In
821 contrast, strong agreement with the item “Despite what he/she did, I want us to have a
822 positive relationship” from the benevolence scale contributed to a higher score.

823 *Single-item measure of forgiveness*

824 We also used a single item, “Have you forgiven the person for the offense?” as a
825 criterion measure of forgiveness. Subjects responded using a slider scale with response
826 options ranging from 0 and 100. Higher scores indicated greater forgiveness.

827

STUDY 4 RESULTS

828 Analyses were conducted using Mplus version 7 (Muthén & Muthén, 1998-2012).
829 See supplementary materials for data and syntax. We report standardized coefficients
830 for all outcomes.

831 *Factor Analysis of the TRIM-18*

832 As in Studies 1-3, we fit three confirmatory correlated models for the TRIM-18
833 using one-factor, two-factor (avoidance-benevolence and revenge), and three-factor
834 (avoidance, benevolence, and revenge) models. Model fit for the factor structures can
835 be seen in Table 2b. Consistent with previous research, we found that the fit of the two-
836 factor model was comparable to that of the three-factor model, which is unsurprising
837 given that the avoidance and benevolence factors in the three-factor model were
838 correlated at $r = .915$. In contrast to what we found with the TRIM-NCO, which
839 consistently favored a model with three correlated factors, we thus found that people's
840 responses to items on the TRIM-18 may be informed by only two underlying constructs.

841 To better understand the implications of the relative interchangeability of the two-
842 factor and three-factor models for an understanding of forgiveness based on the bifactor
843 model, we created two alternative bifactor (S-1) models: For the first model, we created
844 only one specific factor for the revenge items, with the general factor scaled by the
845 avoidance and benevolence items (referred to as a bifactor (2-1) model). For the
846 second model, we created two specific factors—one for the revenge items and another
847 for the benevolence items—with the general factor scaled by the avoidance items (as in
848 Studies 1-3; referred to as a bifactor (3-1) model). As mentioned above, bifactor (S-1)
849 models and an S-correlated factors model (where S represents the number of factors in
850 the comparable models) yield identical model fit under reasonable restrictions (Geiser et

851 al., 2008); therefore, these models were compared for their predictive utility (with a
852 single self-report item of forgiveness as our criterion) against the two correlated factors
853 and three correlated factors models, respectively. Path coefficients can be seen in
854 Table 3b.

855 *Did the Two- and Three-Factor Models Predict Single-Item Forgiveness?*

856 We regressed the single-item measure of forgiveness on each of the factors in
857 the two- (avoidance-benevolence and revenge) and three-factor (avoidance, revenge,
858 and benevolence) models of the TRIM-18. In the two-factor model, the single-item
859 forgiveness measure was predicted by avoidance-benevolence, $b = 0.450$, $SE = 0.074$,
860 $p < .001$, and revenge, $b = 0.232$, $SE = 0.086$, $p < .001$. In the three-factor model, the
861 single-item forgiveness measure was not predicted by avoidance, $b = .072$, $se = .187$, p
862 $= .699$, but was predicted by benevolence, $b = 0.382$, $SE = 0.167$, $p = .022$, and
863 revenge, $b = 0.250$, $SE = 0.091$, $p = .006$.

864 The fact that the avoidance factor in the three-factor model did not predict unique
865 variance in the single-item forgiveness measure lends additional support to our
866 speculation that avoidance and benevolence are indistinguishable in close relationships.
867 In fact, a three-factor model leads to the problematic conclusion that avoidance
868 motivation is unrelated to self-reported forgiveness after accounting for revenge and
869 benevolence motivations, which is not surprising in light of the high collinearity between
870 the avoidance and benevolence factors in the three-factor model.

871 *Did the Bifactor (S-1) Models Predict Single-Item Forgiveness?*

872 As mentioned previously, we specified two bifactor (S-1) models: The first with a
873 specific factor only for the revenge items (bifactor (2-1) model) and the second with

874 specific factors for both the revenge and the benevolence items (bifactor (3-1) model).
875 For both of these models, we regressed the single-item measure of forgiveness on the
876 general forgiveness factor and the specified group factor or factors. In the bifactor (2-1)
877 model (with a specific factor only for the revenge items), scores on the single-item
878 forgiveness measure were significantly predicted by the general factor, $b = .561$, $se =$
879 $.061$, $p < .001$, and the revenge specific factor, $b = -.208$, $se = .073$, $p = .004$. The
880 direction of the relationship between the revenge factor and the single-item measure of
881 forgiveness was troubling, however, because the revenge items here are reverse
882 scored; thus, the negatively signed regression coefficient suggests that disavowal of
883 revenge motivation is associated with lower scores on the single-item measure of
884 forgiveness. In the bifactor (3-1) model (with specific factors for both the revenge items
885 and the benevolence items), single-item forgiveness responses were significantly
886 predicted by the general factor, $b = 0.582$, $SE = 0.060$, $p < .001$, but not the revenge
887 factor, $b = 0.144$, $SE = 0.113$, $p = .202$, or the benevolence factor, $b = -0.044$, $SE =$
888 0.103 , $p = .666$. The significant negative relationship of (disavowals of) revenge
889 motivation with the single-item measure of forgiveness in the bifactor (2-1) model
890 therefore disappeared once the group factor for the benevolence items was also
891 included.

892 Next, we correlated the general factor with a score derived by simply taking the
893 mean of all 18 items from the TRIM-18. The two measures were almost perfectly
894 correlated, $r = .978$, $p < .001$. Unsurprisingly, this simple 18-item composite was also
895 significantly correlated with people's responses to the single-item forgiveness measure,
896 $r = .586$, $p < .001$. As was the case with the TRIM-NCO in Studies 1-3, these results

897 suggest that the general factor can be estimated with high fidelity simply by taking the
898 mean of all of 18 items on the TRIM-18.

899 *Did Item Responses Support a Continuum from Malevolence to Benevolence on the*
900 *General Factor?*

901 As in Studies 1-3, we computed L_{IRF} for each item from the general factor (Ali et
902 al., 2015), which are presented in Table 5b. Again, we found that L_{IRF} values were
903 sorted along the general factor with revenge items on the easier end and a mixture of
904 avoidance and benevolence items on the more difficult end. We also found that L_{IRF}
905 values and L_{IRF} ranks for each item were strongly correlated with those found in Study
906 1 ($r=.861$ and $r=.641$, respectively), Study 2 ($r=.719$ and $r=.741$, respectively), and
907 Study 3 ($r=.906$ and $r=.520$, respectively), providing further evidence for consistency
908 in item difficulties across samples, as well as for the interpretation that the general
909 factor represents a continuum from revenge to avoidance and benevolence.

910 *Is the general factor of forgiveness unique to the TRIM Inventory or does it apply to*
911 *other measures of forgiveness?*

912 Next, we sought to determine whether the single malevolence-benevolence
913 factor that we hypothesize to underlie forgiveness also explains the pattern of item
914 responses on the Decision to Forgive Scale and the Emotional Forgiveness Scale
915 (Hook et al., 2009; Hook et al., 2012). To do so, we fit a bifactor (S-1) model to the 18
916 items from the TRIM Inventory (McCullough et al., 2006), the eight items from Hook and
917 colleague's Decision to Forgive Scale, and the eight items from the Emotional
918 Forgiveness Scale (Hook et al., 2009; Hook et al., 2012). The bifactor (S-1) model
919 included four group factors: two for the revenge and benevolence items from the TRIM

920 Inventory, and two for the items on the Decision to Forgive Scale and the Emotional
921 Forgiveness Scale, respectively. The model exhibited good model fit (see Table 2b),
922 with the general factor explaining common variance across all items in the TRIM
923 Inventory, the Decision to Forgive Scale, and the Emotional Forgiveness Scale.

924 We then regressed the single-item measure of forgiveness on the general factor
925 from this newly created bifactor (S-1) model. As we found with simply modeling
926 responses from the TRIM-18, the general factor remained a strong predictor of the
927 single-item measure of forgiveness, $b = .592$, $se = .059$, $p < .001$. The
928 relationship of the general factor and the single-item measure of forgiveness changed
929 very little in magnitude when we included the Decisional and Emotional Forgiveness
930 scales in the bifactor models ($\Delta b = .031$), indicating that the general factor neither
931 gained nor lost substantial precision by including responses from additional scales. The
932 fact that the general factor estimated exclusively on the basis of the TRIM-18 operates
933 essentially identically to a general factor that is estimated by also including other
934 forgiveness scales stands as evidence that the general factor does in fact reflect
935 subjects' standing on an attitudinal or motivational continuum that exists independently
936 of any specific tool used to measure it.

937 STUDY 4 DISCUSSION

938 In Study 4, we used a sample of non-East Asian U.S. undergraduates to
939 determine if our findings extended to relationships involving close others. We largely
940 replicated the results of Studies 1-3, such that a bifactor (S-1) model of forgiveness
941 provided good model fit across the context of close relationships, and model fit for the
942 bifactor (S-1) remained excellent even as we added items from related scales designed

943 to tap emotional and decisional components of forgiveness. The latter result suggests
944 that the same underlying malevolence-benevolence continuum suffuses the items from
945 a second self-report measure of forgiveness. Moreover, the general factor of the bifactor
946 models predicted a single-item measure of forgiveness, even in models that
947 incorporated self-report measure of forgiveness beyond the TRIM. When we analyzed
948 the three-factor model of forgiveness, the benevolence and revenge factors predicted
949 the single-item measure of forgiveness, but avoidance did not. In Study 4, our analyses
950 of item responses likewise replicated Studies 1 through 3 in suggesting the possibility
951 that a single motivational continuum underlies the forgiveness process, ranging from
952 malevolence to benevolence. In contrast to Studies 1 through 3, however, factor
953 analyses from Study 4 did not reveal significant differences in fit between the two-factor
954 and three-factor models when using the correlated-factors approach. Given that Studies
955 1 through 3 involved strangers interacting for the first time, whereas Study 4 involved
956 close others, this result may have interesting implications for our understanding of how
957 forgiveness operates across different relationship types—a topic to which we return in
958 the General Discussion.

959 STUDY 5

960 Through these experiments, we have sought to replicate our results in different
961 experimental contexts to ensure that our results are generalizable beyond a single
962 sampling procedure or study design. In Study 5, we therefore sought to replicate the
963 results of Study 4—with its emphasis on forgiveness in close relationships—in a sample
964 of Japanese undergraduate students, which we expected would yield consistent results
965 with our U.S. samples, just as Study 3 broadly replicated Studies 1 and 2.

966 STUDY 5 METHOD

967 *Subjects*

968 Subjects were 158 undergraduate students recruited using the psychology
969 subject pool at a university in Japan (51.2% female; Age not collected). East Asian
970 subjects were targeted for recruitment, resulting in an ethnically homogenous sample.
971 Subjects completed the survey in exchange for 700 JPY.

972 *Procedure*

973 The procedure in Study 5 was highly similar to Study 4. Subjects completed
974 back-translated versions of the scales in Study 4 (Ohtsubo et al., 2015). The only
975 procedural deviation was in subject recruitment, as subjects completed the survey in a
976 laboratory setting in order to receive compensation for their participation.

977 STUDY 5 RESULTS

978 Analyses were conducted using Mplus version 7 (Muthén & Muthén, 1998-2012).
979 See supplementary materials for data and syntax. We report standardized coefficients
980 for all outcomes.

981 *Factor Analysis of the TRIM-18*

982 As in Study 4, we fit three confirmatory models for the TRIM-18 using one-factor,
983 two-factor (avoidance-benevolence and revenge), and three-factor (avoidance,
984 benevolence, and revenge) models. Model fit for the factor structures can be seen in
985 Table 2b. Consistent with Study 4, we found that the two-factor and three-factor models
986 fit very comparably, which again is unsurprising in light of the fact that the avoidance
987 and benevolence factors in the three-factor model were correlated at $r = .919$.

988 To be consistent with Study 4, we created two alternative bifactor (S-1) models:
989 For the first model, we created only one specific factor—for the revenge items—with the
990 general factor scaled by the avoidance and benevolence items. For the second model,
991 we created two specific factors—one for the revenge items and another for the
992 benevolence items—with the general factor scaled by the avoidance items (as in
993 Studies 1-3). These models were then compared for their predictive utility (with a single
994 self-report item of forgiveness as our criterion) against the two correlated factors and
995 three correlated factors models, respectively. Path coefficients can be seen in Table 3b.

996 *Did the Two- and Three-Factor Models Predict Single-Item Forgiveness?*

997 We regressed the single-item measure of forgiveness on each of the factors in
998 the two-factor (avoidance-benevolence and revenge) and three-factor (avoidance,
999 revenge, and benevolence) models of the TRIM-18. In the two-factor model, the single-
1000 item forgiveness measure was predicted by avoidance-benevolence, $b = 0.388$, $SE =$
1001 0.073 , $p < .001$, and revenge, $b = 0.276$, $SE = 0.070$, $p < .001$. In the three-factor
1002 model, the single-item forgiveness measure was negatively predicted by avoidance, $b =$
1003 $-.421$, $se = .175$, $p = .016$, and positively by revenge, $b = 0.307$, $SE = 0.067$, $p < .001$,
1004 and benevolence, $b = 0.807$, $SE = 0.179$, $p < .001$.

1005 Although these results did not yield the same problematic patterns we found in
1006 Study 4, we did find that the three-factor model created a negative relationship between
1007 (reverse-scored) avoidance and the single-item measure of forgiveness after accounting
1008 for variance in benevolence. This may be more evidence that, when measured in close
1009 relationships, the benevolence and avoidance items reflect the same underlying

1010 construct, thereby causing problems when trying to identify their unique influence on
1011 forgiveness-relevant constructs.

1012 *Did the Bifactor (S-1) Models Predict Single-Item Forgiveness?*

1013 As mentioned previously, we specified two bifactor (S-1) models: The first with
1014 only one specific factor for the revenge items (bifactor (2-1) model) and the second with
1015 specific factors for both the revenge and the benevolence items (bifactor (3-1) model).
1016 For both of these models, we regressed the single-item measure of forgiveness on the
1017 general forgiveness factor and the respective group factors. In the bifactor (2-1) model
1018 (with a specific factor only for the revenge items), scores on the single-item forgiveness
1019 measure were significantly predicted by the general factor, $b = .532$, $se = .063$, $p < .001$,
1020 and the revenge factor, $b = .243$, $se = .055$, $p < .001$. Unlike in Study 4, the relationship
1021 between the (reverse-scored) revenge factor was positively related to forgiveness, as
1022 one would expect. In the bifactor (3-1) model (with specific factors for both the revenge
1023 items and the benevolence items), single-item forgiveness responses were significantly
1024 predicted by the general factor, $b = 0.539$, $SE = 0.063$, $p < .001$, the revenge factor, $b =$
1025 0.198 , $SE = 0.055$, $p < .001$, and the benevolence factor, $b = 0.178$, $SE = 0.055$, $p =$
1026 $.001$.

1027 Next, we correlated the general factor with a score derived from simply taking the
1028 mean of all 18 items from the TRIM-18. The two measures were highly correlated, $r =$
1029 $.816$, $p < .001$. Unsurprisingly, this simple 18-item composite was also significantly
1030 correlated with people's responses to the single-item forgiveness measure, $r = .589$, $p <$
1031 $.001$. As was the case with the TRIM-NCO in Studies 1-3 and the TRIM-18 in Study 4,

1032 these results suggest that the general factor can be estimated with reasonably high
1033 fidelity simply by taking the mean of all of 18 items on the TRIM-18.

1034 *Did Item Responses Support a Continuum from Malevolence to Benevolence on the*
1035 *General Factor?*

1036 As in Studies 1-3, we computed L_{IRF} for each item (Ali et al., 2015), which are
1037 presented in Table 5b. Again, we found that L_{IRF} values were sorted along the general
1038 factor with revenge items on the easier end and a mixture of avoidance and
1039 benevolence items on the more difficult end. We also found that L_{IRF} values and L_{IRF}
1040 ranks for each item were strongly correlated with those found in Study 1 ($r = .768$ and $r =$
1041 $.525$, respectively), in Study 2 ($r = .730$ and $r = .748$, respectively), in Study 3 ($r = .826$
1042 and $r = .699$, respectively), and in Study 4 ($r = .874$ and $r = .794$, respectively), providing
1043 further evidence for consistency in item difficulties across samples, as well as for the
1044 interpretation that the general factor represents a continuum from malevolence to
1045 benevolence.

1046 *Does the general factor of forgiveness capture dimensionality in other forgiveness*
1047 *measures?*

1048 Next, we sought to determine whether the single malevolence-benevolence
1049 factor that we hypothesize to underlie forgiveness also explains the pattern of item
1050 responses on other forgiveness measures—specifically here the Decision to Forgive
1051 Scale and Emotional Forgiveness Scale (Hook et al., 2009; Hook et al., 2012). As in
1052 Study 4, we fit a bifactor (S-1) model to the 18 items from the TRIM Inventory, the eight
1053 items from Hook and colleagues' Decision to Forgive Scale, and the eight items
1054 Emotional Forgiveness Scale (Hook et al., 2009; Hook et al., 2012). The bifactor (S-1)

1055 model included four group factors: two for the revenge and benevolence items from the
1056 TRIM Inventory, and two for the items on the Decision to Forgive Scale and the
1057 Emotional Forgiveness Scale, respectively. The model exhibited good model fit (see
1058 Table 2b), with the general factor explaining common variance across all items in the
1059 TRIM, Decision to Forgive Scale, and Emotional Forgiveness Scale.

1060 We then regressed the single-item measure of forgiveness on the general factor
1061 from this newly created bifactor (S-1) model. As we found with modeling responses from
1062 the TRIM-18 only, the general factor remained a strong predictor of single-item
1063 forgiveness, $b = .556$, $se = .064$, $p < .001$. As we found in Study 4, the
1064 relationship of the general factor and the single-item measure of forgiveness changed
1065 very little in magnitude when we included the Decisional and Emotional Forgiveness
1066 scales in the bifactor models ($\Delta b = .017$), which stands as further evidence that the
1067 general factor does in fact reflect subjects' standing on an attitudinal or motivational
1068 continuum that exists independently of any specific tool used to measure it.

1069 STUDY 5 DISCUSSION

1070 In Study 5, we sampled from Japanese students at a university in Japan to
1071 replicate the main results we obtained in Study 4 (which were obtained from non-East
1072 Asian students in the U.S.). In concert with the results of Study 4, we found that the
1073 bifactor modeling approach yielded good fit, with a general factor that predicted single-
1074 item forgiveness scores. As in Study 4, we also found that the general factor captured
1075 variance underlying not only the TRIM-18 but also the Decision to Forgive and
1076 Emotional Forgiveness Scales, which suggests that the same underlying malevolence-
1077 benevolence continuum suffuses the items from a second self-report measure of

1078 forgiveness. Together with similar results from Studies 1-3 in which non-close others
1079 were involved, item response analyses suggest that such a motivational or attitudinal
1080 continuum underlies forgiveness in both close and non-close relationships. However, as
1081 in Study 4, factor analyses using the correlated factors approach found no significant
1082 difference in fit between the two-factor and three-factor models. The contrast with the
1083 results of Studies 1 through 3 in this regard provides additional grounds to suspect that
1084 the structure of forgiveness may differ in close vs. non-close relationships. We address
1085 this possibility further in the General Discussion.

1086 GENERAL DISCUSSION

1087 Forgiveness has long been conceptualized as a process of psychological change
1088 regarding a harmdoer, but change in what? Through the years, scholars have proposed
1089 a variety of theoretical models of the psychological changes that constitute forgiveness
1090 (e.g., McCullough, Fincham, & Tsang, 2002; Subkoviak et al., 1995), but the tools they
1091 have developed to measure forgiveness have not always succeeded in capturing those
1092 complexities and nuances. Generally, the constructs defined by theory often appear to
1093 be more complex than the underlying structure of the psychometric instruments
1094 designed to capture those theoretical constructs. Here, with the goal of obtaining a
1095 better theoretical understanding of the underlying structure of the psychological
1096 phenomena that change when people forgive, we tested a variety of modeling
1097 approaches for a commonly used measure of forgiveness—the Transgression-Related
1098 Interpersonal Motivations (TRIM) Inventory (McCullough et al., 2006)—with separate
1099 variants of the scale for close and non-close relationships. In doing so, we specifically
1100 sought to determine whether a bifactor modeling approach, which specifies forgiveness

1101 as a single underlying continuum that ranges from malevolence to benevolence, along
1102 with other potentially substantive item-specific factors, would provide additional clarity
1103 about the psychological dimensions that underlie forgiveness.

1104 In three experiments involving staged transgressions between non-close others,
1105 we found that the bifactor model of responses on the newly developed TRIM-NCO
1106 yielded substantially better model-data fit than did more traditional single-factor and two-
1107 factor confirmatory models. Although the bifactor model and three-factor model fit the
1108 data equally well (which is true by definition, under reasonable constraints; Geiser et al.,
1109 2008), none of the factors in the three-factor model stood out as robust predictors of any
1110 behavioral measure of forgiveness, whereas the general factor of the bifactor model
1111 predicted nearly every behavioral outcome we measured. Indeed, every criterion
1112 variable that was correlated with at least one factor from the three-factor model was
1113 also significantly correlated with the general factor in the bi-factor model—and in the
1114 theoretically expected direction (which was not always the case with the factors derived
1115 from the three-factor model).

1116 Similarly, in two additional studies involving recalled transgressions between
1117 close others, bifactor models of responses to the original TRIM-18 fit the data well, and
1118 the general factor of those bifactor models consistently predicted a single-item self-
1119 report measure of forgiveness—even when we added other self-report measures of
1120 forgiveness (beyond the TRIM-18) into the model. Critically, the general factor that we
1121 obtained from a bifactor model of the TRIM-18 was uncovered with near-perfect
1122 precision from a bifactor model that also included items from two independently
1123 developed measures of forgiveness (Hook et al., 2009, 2012). Across all five studies,

1124 the reliably good fit of the bifactor model, together with the fact that its general factor
1125 consistently predicted relevant behavioral or self-reported criterion measures, strongly
1126 align with an understanding of forgiveness as a general process of change along a
1127 single attitudinal or motivational continuum, perhaps along with one or more group
1128 factors that reflect method variance or some other substantive source of forgiveness-
1129 relevant variance (for example, personality-based response sets that reflect individual
1130 differences in aversion to harming strangers). The psychological meanings and
1131 theoretical importance of the group factors was rendered somewhat uncertain by their
1132 unreliable associations with the various criterion variables with which we sought to
1133 correlate them (for a clear discussion of G and S factors and their meanings in
1134 regressions, see Heinrich, Zagorscak, Eid, & Knaevelsrud, 2018).

1135 Our hypothesis that the trait underlying the measures of forgiveness we
1136 examined here reflects an attitudinal continuum running from malevolence to
1137 benevolence is reinforced by the correlations of the general factor recovered here with
1138 subjects' scores on four different laboratory experimental economics games, including
1139 games that reflect both punitiveness and cooperativeness. Other work has shown that
1140 the variance shared among people's scores in the cooperative games appears to be
1141 caused by a common tendency to cooperate that manifests itself both in the lab and in
1142 real-life social interactions (McAuliffe et al., 2018; Peysakhovich et al., 2014; Wilhelm,
1143 Kaltwasser, & Hildebrandt, 2018), so we are inclined to conclude that the general factor
1144 we found here reflects variation in a single broad motivational or attitudinal construct
1145 that, on the positive end of the continuum, manifests itself through cooperative behavior
1146 in daily life. This conclusion is strengthened considerably by the fact that the item

1147 difficulties of the items on both versions of the TRIM inventory align themselves along a
1148 continuum that ranges from motivation to seek vengeance on one end to motivation to
1149 restore friendly relations on the other. Because a malevolence-benevolence dimension
1150 appears to suffuse interpersonal behavior in general (as exemplified in the Interpersonal
1151 Circumplex model of interpersonal behavior; Gurtman, 2009), it is to some extent
1152 unsurprising to discover that a similar continuum underlies forgiveness, though it has
1153 the potential to be critically important for future theory and research on forgiveness.

1154 *Measuring Forgiveness in Close vs. Non-Close Relationships*

1155 Prior research involving measures of transgression-related interpersonal
1156 motivations has used the TRIM-18 (McCullough et al., 2006) or its earlier 12-item
1157 variant (McCullough et al., 1998), which were designed to gauge the forgiveness-
1158 relevant motivational changes experienced by individuals in close relationships (e.g.,
1159 friendships, family relationships, and mating relationships). In our first three
1160 experiments, we analyzed a new variant of the measure, the TRIM-NCO, which
1161 involved modifying the TRIM-18 to make it applicable to non-close relationships of the
1162 sort that researchers frequently examine in experiments. The novelty of the TRIM-NCO,
1163 with its focus upon interpersonal motivations in new relationships rather than
1164 established ones, raises the concern that the conclusions derived from research that
1165 use it might not generalize across relationship types. We addressed this concern by
1166 reporting the results of two correlational studies that used the standard TRIM-18 to
1167 assess transgression-related motivations in the context of close relationships. For both
1168 close and non-close relationships, we found that the bifactor model showed excellent fit
1169 and that its general factor reliably exhibited predictive utility.

1170 Although the patterns of results involving the bifactor model were largely
1171 consistent across close and non-close relationships, the same cannot be said for results
1172 involving the more standard two-factor and three-factor correlated-factors models. In
1173 studies involving non-close others (Studies 1-3), the two-factor correlated-factor model
1174 (avoidance-benevolence and revenge) fit the data significantly worse than the three-
1175 factor model. But in studies involving close others (Studies 4-5), the two-factor model fit
1176 the data as well as the three-factor model did—a finding consistent with prior work
1177 (McCullough et al., 2007; McCullough et al., 2006). Moreover, in studies involving close
1178 others, both factors of the two-factor model accounted for unique variance in our single-
1179 item criterion measure of forgiveness, whereas the avoidance factor of the three-factor
1180 model did not significantly predict the single item measure after accounting for revenge
1181 and benevolence.

1182 Collectively, therefore, the results of our five studies indicate a discrepancy in the
1183 number and nature of specific factors that underlie forgiveness in close relationships vs.
1184 forgiveness in non-close relationships. The possible causes and implications of this
1185 discrepancy invite consideration. Do avoidance and benevolence motives become
1186 indistinguishable as relationships become close? Does the conflation of avoidance and
1187 benevolence in studies involving close others result from the fact that transgressions
1188 involving close others occurred in the more distant past or because transgressions in
1189 close relationships may already have been resolved at the time of measurement? Is it
1190 because transgressions involving close others are more harmful than those
1191 manufactured in experiments involving strangers? These and other possibilities provide
1192 the basis for future inquiry into the process of forgiveness, although it appears that most

1193 of these questions could be easily elided by concentrating future theoretical efforts on
1194 the general factor (whether estimated explicitly with a bifactor model or, more
1195 approximately, with a sum score of all of the items on the scale).

1196 *Do Specific Factors Reflect Method or Motive?*

1197 We argued from the results of the bifactor (S-1) model of the TRIM Inventory
1198 (and in Studies 4 and 5, both the Decisional and Emotional Forgiveness Scales) that the
1199 general factor reflects a motivational or attitudinal continuum that spans from
1200 malevolence to benevolence. We argued further that the additional group factors reflect
1201 either substantive variation due to theoretically meaningful features of forgiveness or
1202 methodological factors that might best be characterized as nuisance variance. Although
1203 we found some modest correlations between the specific factors and our outcomes,
1204 they were not wholly consistent across experiments, possibly undermining their utility
1205 beyond improving model-data fit. It is also plausible that the specific factors reflect
1206 systematic responses to particular methods (such as positive vs. negative wording) that
1207 distinguish items pertaining to revenge and benevolence, which may also undermine
1208 their predictive utility. In fact, some treatments of a bifactor model are designed to
1209 account for differences in wording (e.g., reverse-scored items) due to inherent
1210 differences in how people respond to items that are worded to be the inverse of their
1211 construct, and scholars have even noted that these applications may be doing most of
1212 their work by accounting for implausible response patterns (Reise, Kim, Mansolf, &
1213 Widaman, 2016). Therefore, we still have some important open questions regarding the
1214 *predictive* utility of modeling specific factors of forgiveness, specifically with regard to
1215 responses on the TRIM Inventory. Again, however, the importance of these findings

1216 seem to us considerably overshadowed by the prominence of the general factor and its
1217 broad conceptual and empirical utility.

1218 One possible limitation of this work is that our first three experiments involved
1219 only a single behavioral measure for each of the constructs representing trust,
1220 benevolence, and punitiveness, which attenuated their reliability. On one hand, this
1221 concern may be somewhat minor for several reasons. First, the iterated prisoner's
1222 dilemma game evinced extremely high internal consistency (McDonald's $\omega_s > .97$; see
1223 supplemental materials) in all three studies, and was positively and significantly
1224 correlated with the general factor (but not the specific factors) in each of them. Second,
1225 previous research has obtained internal consistency estimates for six-item composites
1226 of scores on the Dictator Game and the Trust game of that ranged from .91 to .95 for
1227 the Dictator Game and .95 to .96 for the Trust Game (McAuliffe et al. 2018b). Applying
1228 the Spearman-Brown prophecy formula to the mean of the respective reliability
1229 estimates implies that a one-item measure of the Dictator Game could possess a
1230 reliability as high as .69 and the Trust Game could possess a reliability as high as .76.
1231 These estimates suggest that the one-item constitution of these two measures in our
1232 studies might attenuate true score correlations by as little as $1 - \sqrt{.69} = 17\%$ for the
1233 Dictator Game and $1 - \sqrt{.76} = 13\%$ for the Trust Game. The accuracy of these
1234 prophesied estimates cannot be verified with our data, of course, but they do suggest
1235 that attenuations in the magnitude of the associations we reported here are smaller than
1236 one might imagine. We also remind readers that in Studies 1-3 the general factor was a
1237 robust predictor of most behavioral measures—even the single-item measures—

1238 perhaps because it represents a general propensity to act prosocially, which is in
1239 contrast to the more nuanced interpretations of the specific factors (Geiser et al., 2008).

1240 Even so, our understanding of the relationships between the specific factors of
1241 the TRIM-NCO and behavioral measures of forgiveness could be improved in future
1242 research by assessing forgiveness-relevant cooperative behavior using multiple
1243 indicators per construct (e.g., multiple 'benevolence', 'trust', and 'punitive' behavioral
1244 assessments). We also note that Studies 4 and 5—which used the TRIM-18 and
1245 focused on forgiveness in close relationships—included no behavioral measures, a
1246 limitation that future research might also address.

1247 *Does the Utility of the General Factor Justify a Single-Factor Model?*

1248 Because we were able to gain most of our predictive utility for forgiveness-
1249 relevant outcomes from the general factor of the bifactor (S-1) model, we considered
1250 how well the general factor from this more complex model compares with composite
1251 scores often used in forgiveness research. In each of our studies, we found that the
1252 general factor correlated with an average 'TRIM' score to a very high degree ($r_s > 0.80$;
1253 see supplemental materials), indicating that researchers could retain a great deal of
1254 predictive power by considering only a rudimentary model of forgiveness. Although we
1255 do not recommend this practice, we do think researchers could justify their use of a
1256 simple model of TRIM scores *because* we demonstrated that the general factor of the
1257 bifactor (S-1) is effective at representing the underlying malevolence-benevolence
1258 continuum in which forgiveness evidently takes place. For researchers interested in
1259 applying the bifactor (S-1) model that we advocate here, we make the relevant software
1260 code available in supplemental materials.

1261 *Conclusions*

1262 Although researchers have been studying forgiveness for more than 25 years,
1263 the rich and sometimes multifaceted conceptualizations of forgiveness that have guided
1264 this work have not been tightly moored to the psychometric realities of the tools used to
1265 measure those conceptualizations. The studies presented here strongly support the
1266 hypothesis that forgiveness is a process of attitudinal change, incorporating thoughts,
1267 feelings, and behavioral tendencies, along a dimension that ranges from malevolence to
1268 benevolence. We recovered evidence for this underlying continuum from studies in
1269 laboratory studies as well as in online studies, in the United States as well as in Japan,
1270 with transgressors who are strangers as well as with transgressors who are existing
1271 relationship partners, and with the TRIM Inventory (McCullough et al., 2006) as well as
1272 with the Emotional and Decisional Forgiveness Scales (Hook et al., 2009; Hook et al.,
1273 2012). Perhaps most notably, the underlying malevolence-benevolence dimension we
1274 identified here is behaviorally relevant inasmuch as it is consistently and positively
1275 related to laboratory behaviors that reflect the propensity to cooperate in daily life
1276 (McAuliffe et al., 2018; Peysakovich, Nowak, & Rand, 2014).

1277 Although other factors from the bifactor model apparently underlie the self-report
1278 measures we studied here, their value for inspiring future theoretical and empirical work
1279 remains unclear. For instance, these group factors never uncovered correlates of
1280 forgiveness that the general factor failed to identify, they manifested themselves in
1281 slightly different ways in studies of laboratory transgressions than in studies of real-life
1282 transgressions, and they sometimes yielded nonsensical correlations with the
1283 behavioral targets and single-item measure of forgiveness used here. In contrast, the

1284 general factor explains the huge preponderance of the variance in the individual
1285 questionnaire items, as well as in an external single-item measure of forgiveness. It also
1286 reliably predicts subjects' cooperative behavior, and it never yields nonsensical
1287 correlations with external validation criteria. In other words, the general factor of
1288 responses on self-report forgiveness scales behaves like a measure of forgiveness
1289 should.

1290 Going forward, we believe that substantial theoretical and empirical insights
1291 might be gained by viewing forgiveness as prosocial change along a malevolence-
1292 benevolence attitudinal continuum. In particular, the proposition that forgiveness reflects
1293 attitude change implies that much could be learned about forgiveness by applying the
1294 basic principles that social psychologists have already discovered about attitudes and
1295 how to change them (Albarracin & Shavitt, 2018; Bohner & Dickel, 2011; Dalege et al.,
1296 2016). More generally, we believe and hope that these results demonstrate the promise
1297 of seeking to achieve a closer union between our theories of forgiveness and the
1298 empirical realities of the measures with which we seek to measure the constructs those
1299 theories invoke.

Table 1. Order of events for Studies 1, 2, and 3, including the partner's preprogrammed behavior, the timing of assessments for relationship value and exploitation risk, and the confederate's messages to subjects.

Experiment Stage	Description	
Prisoner's Dilemma Game	Round	Event
	1	Confederate Cooperates
	2-12	Confederate plays generous tit-for-tat <i>Message encouraging continued cooperation</i>
	13-19	Confederate defects <i>Message: Apology Manipulation</i>
	20-21	Confederate Cooperates
	22-29	Confederate plays generous tit-for-tat
Second Economic Game	Trust Game (Studies 1, 2, and 3) or Dictator Game (Study 1) or Pay-to-Punish Game (Study 1)	
Forgiveness	Transgression-Related Interpersonal Motivations	

Note. This table of the experimental protocols includes only the events that pertain to this manuscript. Other measures were taken during the PDG but are not reported here because they are beyond the scope of this manuscript; results using these measures are available in the manuscript written to be a complement to this (Billingsley et al., in prep).

Table 2a. Model fit for different factor structures of the TRIM-NCO.

	χ^2 * (df)	RMSEA [90% CI]	CFI	TLI
Study 1 (U.S. Mechanical Turk)				
One Factor	4493.282 (119)	.190 [.185, .195]	0.921	0.909
Two Factor	2748.931 (118)	.148 [.143, .153]	0.952	0.945
Three Factor	1616.899 (116)	.113 [.108, .118]	0.973	0.968
Bifactor (S-1)	915.451 (108)	.086 [.081, .091]	0.985	0.982
Study 2 (U.S. University/Community)				
One Factor	849.289 (119)	.164 [.154, .175]	0.879	0.862
Two Factor	543.428 (118)	.126 [.115, .137]	0.930	0.919
Three Factor	392.458 (116)	.102 [.091, .113]	0.954	0.946
Bifactor (S-1)	352.619 (108)	.100 [.088, .111]	0.960	0.949
Study 3 (Japanese University)				
One Factor	818.650 (119)	.178 [.166, .189]	0.890	0.875
Two Factor	589.404 (118)	.147 [.135, .158]	0.926	0.915
Three Factor	522.736 (116)	.137 [.125, .149]	0.936	0.925
Bifactor (S-1)	328.437 (108)	.105 [.092, .118]	0.965	0.957

Note: χ^2 * differences cannot be compared directly in a typical χ^2 difference test.

1301

Table 2b. Model fit for different factor structures of the TRIM-18.

	χ^2 * (df)	RMSEA [90% CI]	CFI	TLI
Study 4 (U.S. University)				
One Factor	777.548 (135)	.168 [.157, .180]	0.926	0.916
Two Factor	305.932 (134)	.087 [.074, .100]	0.980	0.977
Three Factor	271.662 (132)	.079 [.066, .093]	0.984	0.981
Bifactor (2-1)	315.680 (130)	.092 [.079, .105]	0.979	0.975
Bifactor (3-1)	261.947 (123)	.082 [.068, .096]	0.984	0.980
Bifactor (w/ DFS and EFS)	821.947 (523)	.058 [.051, .066]	0.977	0.974
Study 5 (Japanese University)				
One Factor	898.322 (135)	.189 [.178, .201]	0.881	0.865
Two Factor	466.095 (134)	.125 [.113, .138]	0.948	0.941
Three Factor	447.372 (132)	.123 [.111, .136]	0.951	0.943
Bifactor (2-1)	466.167 (130)	.128 [.116, .141]	0.947	0.938
Bifactor (3-1)	389.809 (123)	.117 [.104, .130]	0.958	0.948
Bifactor (w/ DFS and EFS)	948.251 (523)	.072 [.064, .079]	0.960	0.955

Note: χ^2 * differences cannot be compared directly in a typical χ^2 difference test.

1302

1303

Table 3a. Path coefficients for economic game decisions predicted by three-factor and bifactor (S-1) models of the TRIM-NCO in Studies 1-3.

		Three-Factor Model			Bifactor (S-1) Model		
Outcome		Avoid	Revenge	Benevolence	General Factor	Revenge-Specific	Benevolence-Specific
Study 1	PDG	-.409 (.097)***	-.109 (.073)	-.011 (.131)	-.493 (.035)***	-.080 (.056)	-.011 (.058)
	TG	-.048 (.130)	-.121 (.107)	.479 (.181)**	.293 (.051)***	-.110 (.083)	.253 (.084)**
	DG	.240 (.186)	.157 (.124)	.004 (.241)	.350 (.049)***	.161 (.096)	-.061 (.085)
	PTPG	-.074 (.141)	-.462 (.096)***	.157 (.176)	-.218 (.056)***	-.343 (.074)***	.027 (.078)
Study 2	PDG	-.250 (.133)	-.154 (.122)	-.067 (.134)	-.403 (.062)***	-.101 (.092)	-.056 (.094)
	TG	.187 (.127)	-.287 (.107)**	.358 (.117)**	.263 (.064)***	-.235 (.078)**	.224 (.081)**
Study 3	PDG	-.099 (.178)	.178 (.155)	-.441 (.178)*	-.379 (.066)***	.139 (.120)	-.148 (.097)
	TG	-.030 (.158)	.102 (.128)	-.007 (.159)	.030 (.070)	.077 (.100)	-.002 (.091)

Notes: Estimates are reported as standardized with standard errors. Asterisks indicate p-value ranges: * $p < .05$, ** $p < .01$, *** $p < .001$

1304

1305

Table 3b. Path coefficients for single-item forgiveness predicted by three-factor and bifactor (S-1) models of the TRIM-18 in Studies 4 and 5.

	Three-Factor Model			Bifactor (3-1) Model		
	Avoid	Revenge	Benevolence	General Factor	Revenge-Specific	Benevolence-Specific
Study 4	.072 (.187)	.250 (.091)**	.382 (.167)*	.582 (.060)***	.144 (.113)	-.044 (.103)
Study 5	-.421 (.175)*	.307 (.067)***	.807 (.179)***	.539 (.063)***	.198 (.055)***	.178 (.055)**
	Two-Factor Model		Bifactor (2-1) Model			
	Avoidance-Benevolence	Revenge	General Factor	Revenge-Specific		
Study 4	.450 (.074)***	.232 (.086)**	.561 (.061)***	-.208 (.073)**		
Study 5	.388 (.073)***	.276 (.070)***	.532 (.063)***	.243 (.055)***		

Notes: Estimates are reported as standardized with standard errors. Asterisks indicate p-value ranges: * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 4a. Correlations between simple composite of 17 items from TRIM-NCO, the bifactor (S-1) model, and behavioral outcomes.

	Study 1			Study 2			Study 3		
	<i>r</i>	<i>se</i>	<i>p</i>	<i>r</i>	<i>se</i>	<i>p</i>	<i>r</i>	<i>se</i>	<i>p</i>
General	0.946	0.004	<.001	0.937	0.009	<.001	0.940	0.012	<.001
Revenge-Specific	0.279	0.011	<.001	0.313	0.026	<.001	0.248	0.032	<.001
Benevolence-Specific	0.275	0.009	<.001	0.365	0.023	<.001	0.306	0.026	<.001
PDG	-0.476	0.033	<.001	-0.414	0.057	<.001	-0.318	0.069	<.001
TG	0.313	0.052	<.001	0.232	0.062	<.001	0.043	0.073	0.557
DG	0.360	0.050	<.001	-	-	-	-	-	-
PTP	-0.300	0.051	<.001	-	-	-	-	-	-

1307

1308

Table 4b. Correlations between simple composite of 18 items from TRIM-18, the bifactor (S-1) model, and single-item forgiveness.

	Study 4			Study 5		
	<i>r</i>	<i>se</i>	<i>p</i>	<i>r</i>	<i>se</i>	<i>p</i>
General	0.978	0.009	<.001	0.953	0.011	<.001
Revenge-Specific	0.092	0.029	.001	0.189	0.025	<.001
Benevolence-Specific	0.033	0.024	.174	-0.003	0.027	.903
Single-Item	0.586	0.051	<.001	0.589	0.052	<.001

1309

Table 5a. Raw means for item responses and item difficulties based on the Location Index from the Item Response Function in the three Experiments. Each experiment is sorted by item difficulty (easiest to hardest).

Study 1				Study 2				Study 3			
<i>Item</i>	Difficulty	Mean	Scale	<i>Item</i>	Difficulty	Mean	Scale	<i>Item</i>	Difficulty	Mean	Scale
17	-3.91	4.35	REV	17	-2.75	3.95	REV	17	-4.85	4.39	REV
4	-3.09	4.49	REV	1	-1.78	3.9	REV	13	-4.16	4.34	REV
1	-2.55	4.2	REV	4	-1.67	4.1	REV	1	-3.07	4.35	REV
13	-2.33	4.15	REV	8	-1.16	3.71	BEN	4	-2.70	4.49	REV
16	-1.40	3.86	BEN	2	-1.14	3.66	AVO	16	-1.08	3.69	BEN
8	-0.92	3.53	BEN	18	-1.05	3.56	AVO	18	-0.97	3.57	AVO
3	-0.90	3.56	BEN	11	-0.99	3.57	AVO	10	-0.90	3.52	BEN
14	-0.81	3.67	BEN	5	-0.91	3.53	AVO	14	-0.88	3.53	BEN
10	-0.71	3.41	BEN	13	-0.80	3.62	REV	8	-0.86	3.52	BEN
18	-0.65	3.31	AVO	15	-0.75	3.49	AVO	5	-0.81	3.42	AVO
11	-0.59	3.3	AVO	3	-0.75	3.59	BEN	11	-0.80	3.47	AVO
5	-0.58	3.35	AVO	10	-0.64	3.51	BEN	2	-0.77	3.42	AVO
6	-0.56	3.37	BEN	14	-0.63	3.53	BEN	6	-0.54	3.46	BEN
2	-0.55	3.33	AVO	16	-0.59	3.48	BEN	15	-0.50	3.32	AVO
15	-0.19	3.12	AVO	6	-0.54	3.45	BEN	7	-0.50	3.28	AVO
12	-0.07	3.03	AVO	12	-0.41	3.31	AVO	3	-0.46	3.25	BEN
7	0.37	2.74	AVO	7	-0.05	3.17	AVO	12	-0.02	2.99	AVO

Note: Item prompts can be referenced using Appendix A. Difficulty values represent the Location Index based on the Item Response Function (L_{IRF} ; see supplemental materials for raw values and computation). Mean values represents the average observed item response, scored so that larger values indicate greater forgiveness. Item Scales represent *a priori* conceptualizations of the items: REV = Revenge; AVO = Avoidance; BEN = Benevolence.

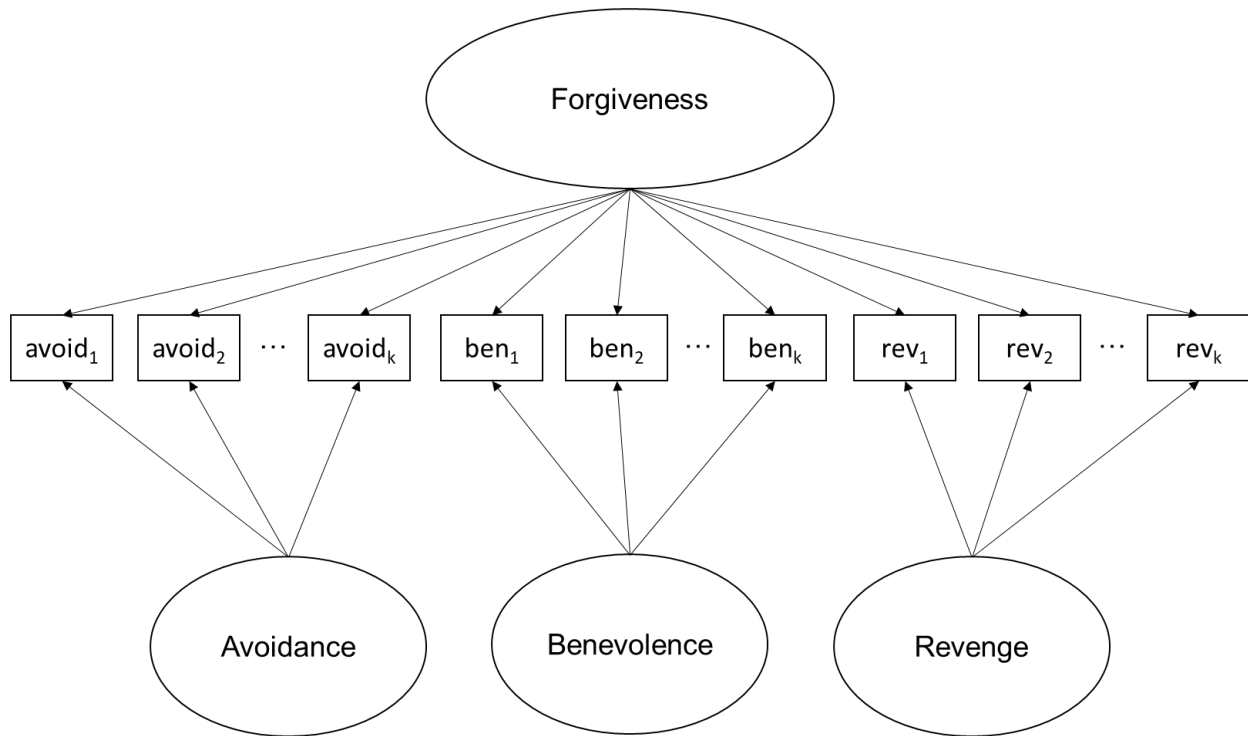
1310

1311

Table 5b. Raw means for item responses and item difficulties based on the Location Index from the Item Response Function in the two studies. Items within each study are sorted by difficulty (easiest to hardest).

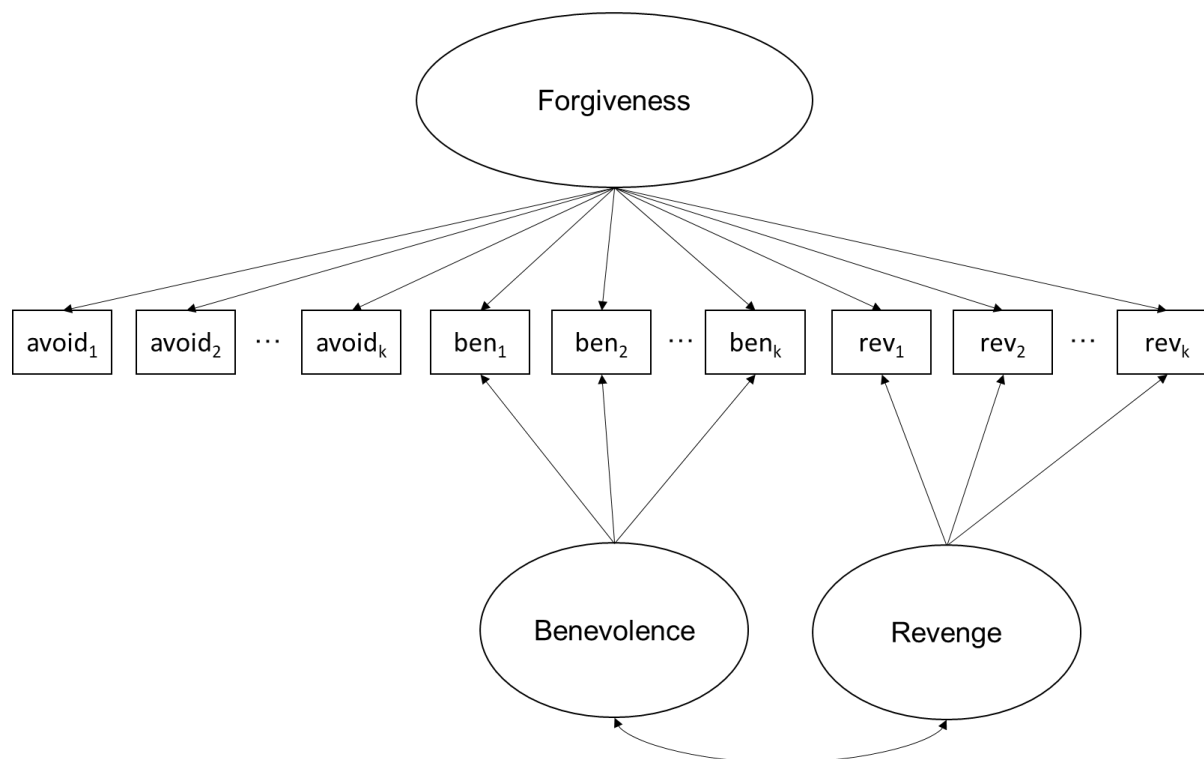
Study 4				Study 5			
<i>Item</i>	Difficulty	Mean	Scale	<i>Item</i>	Difficulty	Mean	Scale
13	-1.918		REV	1	-1.565		REV
17	-1.798		REV	17	-1.245		REV
1	-1.611		REV	13	-1.239		REV
4	-1.49		REV	4	-1.028		REV
9	-1.103		REV	9	-0.981		REV
3	-0.91		BEN	15	-0.89		AVO
15	-0.633		AVO	18	-0.7		AVO
8	-0.508		BEN	10	-0.532		BEN
5	-0.505		AVO	11	-0.514		AVO
11	-0.504		AVO	5	-0.504		AVO
18	-0.424		AVO	2	-0.451		AVO
10	-0.384		BEN	8	-0.414		BEN
2	-0.328		AVO	12	-0.38		AVO
6	-0.281		BEN	14	-0.308		BEN
14	-0.257		BEN	6	-0.303		BEN
12	-0.187		AVO	16	-0.255		BEN
7	-0.123		AVO	3	-0.214		BEN
16	-0.055		BEN	7	-0.059		AVO

Note: Item prompts can be referenced using Appendix B. Difficulty values represent the Location Index based on the Item Response Function (LIIRF; see supplemental materials for raw values and computation). Mean values represents the average observed item response, scored so that larger values indicate greater forgiveness. Item Scales represent *a priori* conceptualizations of the items: REV = Revenge; AVO = Avoidance; BEN = Benevolence. Item 9 was removed from the three experiments due to poor fit; however, it was retained in the two non-experimental studies.



1313
 1314
 1315
 1316
 1317
 1318

Figure 1. Traditional bifactor model applied to item responses on the TRIM-18, which measures three specific motives (avoidance, benevolence, and revenge) to construct a general forgiveness measure.



1319
 1320
 1321
 1322
 1323

Figure 2. A bifactor (S-1) model applied to the TRIM-18. Specific domains are modeled separately from the general factor, with the exception of a reference domain, which is used to define the scale of the general factor.

1324 Appendix A. Items from the Transgression Related Interpersonal Motivations
 1325 questionnaire for Non-Close Others.

1326

1327 *For the following questions, please indicate your current thoughts and feelings about*
 1328 *[Target] using the scale below. Even though you will never encounter [Target] again, we*
 1329 *are interested in how you think you would respond if you were to encounter [Target] in*
 1330 *your daily life.*

1331

1332 1 = strongly disagree

1333 2 = mildly disagree

1334 3 = agree and disagree equally

1335 4 = mildly agree

1336 5 = strongly agree

1337 6 = I prefer not to answer

1338

1339 1. I would want to make him/her pay for treating me badly today.

1340 2. I would try to keep as much distance between the two of us as possible.

1341 3. I would have good will for him/her.

1342 4. I would hope for something bad to happen to him/her.

1343 5. I would have nothing to do with him/her.

1344 6. I would try to put aside any reservations I had in order to develop a good relationship
 1345 with him/her.

1346 7. I would not trust him/her.

1347 8. I would be willing to work toward a positive relationship with him/her.

1348 9. I would want to see him/her get what he/she deserves.

1349 10. I would act warmly towards him/her.

1350 11. I would avoid contact with him/her.

1351 12. I would be very happy to interact with him/her.

1352 13. I would want to get even with him/her.

1353 14. I would try to give up negative feelings toward him/her.

1354 15. I would avoid working with him/her.

1355 16. I would be willing to let go of my anger towards him/her.

1356 17. I would want to seek revenge.

1357 18. I would try to avoid him/her.

1358 *Note: The 9th item exhibited low factor loadings across all modeling techniques and*
 1359 *experiments; therefore, we suggest that future research omit this item to create a scale*
 1360 *of 17 items.*

1361

1362

1363 Appendix B. Items from the Transgression Related Interpersonal Motivations
 1364 questionnaire for Close Others (TRIM-18).

1365
 1366 *For the following questions, please indicate your current thoughts and feelings about the*
 1367 *person who hurt you; that is, we want to know how you feel about that person **right***
 1368 ***now**. Next to each item, circle the number that best describes your current thoughts and*
 1369 *feelings.*

1370
 1371 1 = strongly disagree

1372 2 = disagree

1373 3 = neutral

1374 4 = agree

1375 5 = strongly agree

1376

1377 1. I'll make him or her pay.

1378 2. I am trying to keep as much distance between us as possible.

1379 3. Even though his/her actions have hurt me, I have good will for him/her.

1380 4. I wish that something bad would happen to him/her.

1381 5. I am living as if he/she doesn't exist, isn't around.

1382 6. I want us to bury the hatchet and move forward with our relationship.

1383 7. I don't trust him/her.

1384 8. Despite what he/she did, I want us to have a positive relationship.

1385 9. I want him/her to get what he/she deserves.

1386 10. I am finding it difficult to act warmly towards him/her.

1387 11. I am avoiding him/her.

1388 12. Although he/she hurt me, I am putting the hurt aside so we can resume our
 1389 relationship.

1390 13. I'm going to get even.

1391 14. I have given up my hurt and resentment.

1392 15. I cut off the relationship with him/her.

1393 16. I have released my anger so I can work on restoring our relationship to health.

1394 17. I want to see him or her hurt and miserable.

1395 18. I withdraw from him/her.

1396

1397 Appendix C. Items from the Decisional Forgiveness Scale (DFS).

1398

1399 *Think of your current intentions toward the person who hurt you. Indicate the degree to*
1400 *which you agree or disagree with the following statements.*

1401

1402 1 = strongly disagree

1403 2 = disagree

1404 3 = neutral

1405 4 = agree

1406 5 = strongly agree

1407

1408 1. I intend to try to hurt him or her in the same way he or she hurt me.

1409 2. I will not try to help him or her if he or she needs something.

1410 3. If I see him or her, I will act friendly.

1411 4. I will try to get back at him or her.

1412 5. I will try to act toward him or her in the same way I did before he or she hurt me.

1413 6. If there is an opportunity to get back at him or her, I will take it.

1414 7. I will not talk with him or her.

1415 8. I will not seek revenge upon him or her.

1416

1417 Appendix D. Items from the Emotional Forgiveness Scale (EFS).

1418

1419 *Think of your current emotions toward the person who hurt you. Indicate the degree to*
1420 *which you agree or disagree with the following statements.*

1421

1422 1 = strongly disagree

1423 2 = disagree

1424 3 = neutral

1425 4 = agree

1426 5 = strongly agree

1427

1428 1. I care about him or her.

1429 2. I no longer feel upset when I think of him or her.

1430 3. I'm bitter about what he or she did to me.

1431 4. I feel sympathy toward him or her.

1432 5. I'm mad about what happened.

1433 6. I like him or her.

1434 7. I resent what he or she did to me.

1435 8. I feel love toward him or her.

1436

1437

1438 References

- 1439 Albarracín, D., & Shavitt, S. (2018). Attitudes and attitude change. *Annual Review of*
1440 *Psychology*, 69, 299-327.
- 1441 Ali, U. S., Chang, H. H., & Anderson, C. (2015). Location Indices for Ordinal Polytomous
1442 Items Based on Item Response Theory. *ETS Research Report Series*, 2015(2),
1443 1-13.
- 1444 Aronson, E., Carlsmith, J. M., & Ellsworth, P. C. (1990). *Methods of research in social*
1445 *psychology*. McGraw-Hill New York.
- 1446 Axelrod, R. (1980). More effective choice in the prisoner's dilemma. *Journal of Conflict*
1447 *Resolution*, 24(3), 379-403.
- 1448 Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history.
1449 *Games and economic behavior*, 10(1), 122-142.
- 1450 Billingsley, J., Forster, D. E., Russell, V. M., Burnette, J. L., Ohtsubo, Y., Lieberman, D.,
1451 & McCullough, M. E. (in prep). Apologies promote forgiveness and restore
1452 cooperation by increasing perceived relationship value.
- 1453 Billingsley, J., & Losin, E. A. (2017). The neural systems of forgiveness: an evolutionary
1454 psychological perspective. *Frontiers in Psychology*, 8, 737.
- 1455 Bohner, G., & Dickel, N. (2011). Attitudes and attitude change. *Annual Review of*
1456 *Psychology*, 62, 391-417.
- 1457 Bono, G., McCullough, M. E., & Root, L. M. (2008). Forgiveness, feeling connected to
1458 others, and well-being: Two longitudinal studies. *Personality and Social*
1459 *Psychology Bulletin*, 34(2), 182-195.
- 1460 Buchanan, E. M., & Scofield, J. E. J. B. r. m. (2018). Methods to detect low quality data
1461 and its implication for psychological research. 1-11.
- 1462 Buhrmester, M., Kwang, T., & Gosling, S. D. J. P. o. p. s. (2011). Amazon's Mechanical
1463 Turk: A new source of inexpensive, yet high-quality, data? , 6(1), 3-5.
- 1464 Carlisle, R. D., Tsang, J.-A., Ahmad, N. Y., Worthington Jr, E. L., Witvliet, C. V., &
1465 Wade, N. (2012). Do actions speak louder than words? Differential effects of
1466 apology and restitution on behavioral and self-report measures of forgiveness.
1467 *The Journal of Positive Psychology*, 7(4), 294-305.
- 1468 Carmody, P., Gordon, K. J. P., & Differences, I. (2011). Offender variables: Unique
1469 predictors of benevolence, avoidance, and revenge? , 50(7), 1012-1017.
- 1470 Dalege, J., Borsboom, D., van Harreveld, F., van den Berg, H., Conner, M., & van der
1471 Maas, H. L. (2016). Toward a formalized account of attitudes: The Causal
1472 Attitude Network (CAN) model. *Psychological Review*, 123(1), 2.
- 1473 De Waal, F., & Pokorny, J. (2005). Primate questions about the art and science of
1474 forgiving. In E. L. Worthington Jr (Ed.), *Handbook of Forgiveness* (pp. 17-32).
1475 New York, NY: Routledge.
- 1476 Dorn, K., Hook, J. N., Davis, D. E., Van Tongeren, D. R., & Worthington Jr, E. L. (2014).
1477 Behavioral methods of assessing forgiveness. *The Journal of Positive*
1478 *Psychology*, 9(1), 75-80.
- 1479 Eid, M., Geiser, C., Koch, T., & Heene, M. (2017). Anomalous results in G-factor
1480 models: Explanations and alternatives. *Psychological Methods*, 22(3), 541.
- 1481 Elliot, A. J., & Thrash, T. M. (2002). Approach-avoidance motivation in personality:
1482 approach and avoidance temperaments and goals. *Journal of personality and*
1483 *social psychology*, 82(5), 804.

- 1484 Exline, J. J., Baumeister, R. F., Bushman, B. J., Campbell, W. K., & Finkel, E. J. (2004).
 1485 Too proud to let go: narcissistic entitlement as a barrier to forgiveness. *Journal of*
 1486 *personality and social psychology*, 87(6), 894.
- 1487 Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution*
 1488 *and Human Behavior*, 25(2), 63-87.
- 1489 Forsythe, R., Horowitz, J. L., Savin, N. E., & Sefton, M. (1994). Fairness in simple
 1490 bargaining experiments. *Games and economic behavior*, 6(3), 347-369.
- 1491 Geiser, C., Eid, M., & Nussbeck, F. W. (2008). On the meaning of the latent variables in
 1492 the CT-C (M-1) model: A comment on Maydeu-Olivares and Coffman (2006).
 1493 *Psychological Methods*, 13(1), 49.
- 1494 Gupta, S. K. (2011). Intention-to-treat concept: a review. *Perspectives in Clinical*
 1495 *Research*, 2(3), 109.
- 1496 Gurtman, M. B. (2009). Exploring personality with the interpersonal circumplex. *Social*
 1497 *and Personality Psychology Compass*, 3(4), 601-619.
- 1498 Harper, Q., Worthington Jr, E. L., Griffin, B. J., Lavelock, C. R., Hook, J. N., Vrana, S.
 1499 R., & Greer, C. L. J. J. o. C. P. (2014). Efficacy of a workbook to promote
 1500 forgiveness: A randomized controlled trial with university students. 70(12), 1158-
 1501 1169.
- 1502 Heinrich, M., Zagorscak, P., Eid, M., & Knaevelsrud, C. (2018). Giving G a Meaning: An
 1503 Application of the Bifactor-(S-1) Approach to Realize a More Symptom-Oriented
 1504 Modeling of the Beck Depression Inventory–II. *Assessment*, 1073191118803738.
- 1505 Hendriks, A. (2012). SoPHIE—Software platform for human interaction experiments.
 1506 *University of Osnabrück, Osnabrück*.
- 1507 Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world?
 1508 *Behavioral and Brain Sciences*, 33(2-3), 61-83.
- 1509 Holzinger, K. J., & Swineford, F. (1937). The bi-factor method. *Psychometrika*, 2(1), 41-
 1510 54.
- 1511 Hook, J. N., Worthington Jr, E. L., & Utsey, S. O. (2009). Collectivism, forgiveness, and
 1512 social harmony. *The Counseling Psychologist*, 37(6), 821-847.
- 1513 Hook, J. N., Worthington Jr, E. L., Utsey, S. O., Davis, D. E., & Burnette, J. L. (2012).
 1514 Collectivistic self-construal and forgiveness. *Counseling and Values*, 57(1), 109-
 1515 124.
- 1516 Kan, I. P., & Drummey, A. B. J. C. i. H. B. (2018). Do imposters threaten data quality?
 1517 An examination of worker misrepresentation and downstream consequences in
 1518 Amazon's Mechanical Turk workforce. 83, 243-253.
- 1519 Kashima, Y., Yamaguchi, S., Kim, U., Choi, S.-C., Gelfand, M. J., & Yuki, M. (1995).
 1520 Culture, gender, and self: a perspective from individualism-collectivism research.
 1521 *Journal of personality and social psychology*, 69(5), 925.
- 1522 McAuliffe, W. H., Forster, D. E., Pedersen, E. J., & McCullough, M. E. (2018). Does
 1523 Cooperation in the Laboratory Reflect the Operation of a Broad Trait? *European*
 1524 *Journal of Personality*.
- 1525 McCullough, M. E., Bono, G., & Root, L. M. (2007). Rumination, emotion, and
 1526 forgiveness: three longitudinal studies. *Journal of personality and social*
 1527 *psychology*, 92(3), 490.

- 1528 McCullough, M. E., Fincham, F. D., & Tsang, J.-A. (2003). Forgiveness, forbearance,
1529 and time: the temporal unfolding of transgression-related interpersonal
1530 motivations. *Journal of personality and social psychology*, *84*(3), 540.
- 1531 McCullough, M. E., & Hoyt, W. T. (2002). Transgression-related motivational
1532 dispositions: Personality substrates of forgiveness and their links to the Big Five.
1533 *Personality and Social Psychology Bulletin*, *28*(11), 1556-1573.
- 1534 McCullough, M. E., Kurzban, R., & Tabak, B. A. (2013). Putting revenge and
1535 forgiveness in an evolutionary context. *Behavioral and Brain Sciences*, *36*(1), 41-
1536 58.
- 1537 McCullough, M. E., Luna, L. R., Berry, J. W., Tabak, B. A., & Bono, G. (2010). On the
1538 form and function of forgiving: modeling the time-forgiveness relationship and
1539 testing the valuable relationships hypothesis. *Emotion*, *10*(3), 358.
- 1540 McCullough, M. E., Pargament, K. I., & Thoresen, C. E. (2001). *Forgiveness: Theory,*
1541 *research, and practice*: Guilford Press.
- 1542 McCullough, M. E., Pedersen, E. J., Tabak, B. A., & Carter, E. C. (2014). Conciliatory
1543 gestures promote forgiveness and reduce anger in humans. *Proceedings of the*
1544 *National Academy of Sciences*, *111*(30), 11211-11216.
- 1545 McCullough, M. E., Rachal, K. C., Sandage, S. J., Worthington Jr, E. L., Brown, S. W.,
1546 & Hight, T. L. (1998). Interpersonal forgiving in close relationships: II. Theoretical
1547 elaboration and measurement. *Journal of personality and social psychology*,
1548 *75*(6), 1586.
- 1549 McCullough, M. E., Root, L. M., & Cohen, A. D. (2006). Writing about the benefits of an
1550 interpersonal transgression facilitates forgiveness. *Journal of consulting and*
1551 *clinical psychology*, *74*(5), 887.
- 1552 Muthén, B. O., & Muthén, L. K. (1998-2012). Mplus User's Guide. Seventh Edition. Los
1553 Angeles, CA: Muthén & Muthén.
- 1554 Ohtsubo, Y., Matsunaga, M., Tanaka, H., Suzuki, K., Kobayashi, F., Shibata, E., . . .
1555 Ohira, H. (2018). Costly apologies communicate conciliatory intention: an fMRI
1556 study on forgiveness in response to costly apologies. *Evolution and Human*
1557 *Behavior*, *39*(2), 249-256.
- 1558 Ohtsubo, Y., & Watanabe, E. (2009). Do sincere apologies need to be costly? Test of a
1559 costly signaling model of apology. *Evolution and Human Behavior*, *30*(2), 114-
1560 123.
- 1561 Ohtsubo, Y., Watanabe, E., Kim, J., Kulas, J. T., Muluk, H., Nazar, G., . . . Zhang, J.
1562 (2012). Are costly apologies universally perceived as being sincere? A test of the
1563 costly apology-perceived sincerity relationship in seven countries. *Journal of*
1564 *Evolutionary Psychology*, *10*(4), 187-204.
- 1565 Ohtsubo, Y., Yamaura, K., & Yagi, A. J. J. J. o. S. P. (2015). Development of Japanese
1566 measures of reconciliatory tendencies. *31*(2), 135-142.
- 1567 Orathinkal, J., Vansteenwegen, A., Enright, R. D., & Stroobants, R. (2007). Further
1568 validation of the Dutch version of the Enright Forgiveness Inventory. *Community*
1569 *Mental Health Journal*, *43*(2), 109-128.
- 1570 Peysakhovich, A., Nowak, M. A., & Rand, D. G. (2014). Humans display a 'cooperative
1571 phenotype' that is domain general and temporally stable. *Nature*
1572 *Communications*, *5*, 4939.

- 1573 Psychology Software Tools, I. (2012). E-Prime (Version 2.0). Retrieved from
 1574 <http://www.pstnet.com>.
- 1575 Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate*
 1576 *Behavioral Resesarch*, 47(5), 667-696.
- 1577 Reise, S. P., Kim, D. S., Mansolf, M., & Widaman, K. F. (2016). Is the bifactor model a
 1578 better model or is it just better at modeling implausible responses? Application of
 1579 iteratively reweighted least squares to the Rosenberg Self-Esteem Scale.
 1580 *Multivariate Behavioral Resesarch*, 51(6), 818-838.
- 1581 Rusbult, C. E., Hannon, P. A., Stocker, S. L., & Finkel, E. J. (2005). Forgiveness and
 1582 relational repair. In E. L. Worthington Jr (Ed.), *Handbook of Forgiveness* (pp.
 1583 185-205). New York, NY: Routledge.
- 1584 Rye, M. S., Loiacono, D. M., Folck, C. D., Olszewski, B. T., Heim, T. A., & Madia, B. P.
 1585 (2001). Evaluation of the psychometric properties of two forgiveness scales.
 1586 *Current Psychology*, 20(3), 260-277.
- 1587 Schug, J., Yuki, M., Horikawa, H., & Takemura, K. (2009). Similarity attraction and
 1588 actually selecting similar others: How cross-societal differences in relational
 1589 mobility affect interpersonal similarity in Japan and the USA. *Asian Journal of*
 1590 *Social Psychology*, 12(2), 95-103.
- 1591 Subkoviak, M. J., Enright, R. D., Wu, C.-R., Gassin, E. A., Freedman, S., Olson, L. M.,
 1592 & Sarinopoulos, I. (1995). Measuring interpersonal forgiveness in late
 1593 adolescence and middle adulthood. *Journal of Adolescence*, 18(6), 641-655.
- 1594 Tabak, B. A., McCullough, M. E., Luna, L. R., Bono, G., & Berry, J. W. (2012).
 1595 Conciliatory gestures facilitate forgiveness and feelings of friendship by making
 1596 transgressors appear more agreeable. *Journal of personality*, 80(2), 503-536.
- 1597 Tsang, J. a., McCullough, M. E., & Fincham, F. D. (2006). The longitudinal association
 1598 between forgiveness and relationship closeness and commitment. *Journal of*
 1599 *Clinical Psychology*, 25(4), 448-472.
- 1600 Wilhelm, O., Kaltwasser, L., & Hildebrandt, A. (2018). Will the real factors of prosociality
 1601 please stand up? a comment on böckler, tusche, and singer (2016). *Social*
 1602 *Psychological and Personality Science*, 9(4), 493-499.
- 1603 Worthington, E. L., Lavelock, C., vanOyen Witvliet, C., Rye, M. S., Tsang, J.-A., &
 1604 Toussaint, L. (2015). Measures of forgiveness: Self-report, physiological,
 1605 chemical, and behavioral indicators. In *Measures of personality and social*
 1606 *psychological constructs* (pp. 474-502): Elsevier.
- 1607 Yamagishi, T., & Yamagishi, M. (1994). Trust and commitment in the United States and
 1608 Japan. *Motivation and Emotion*, 18(2), 129-166.

1609