Dispositional Forgivingness: Development and Construct Validity of the Transgression Narrative Test of Forgivingness (TNTF)

Jack W. Berry
Everett L. Worthington, Jr.
Virginia Commonwealth University

Les Parrott, III
Seattle Pacific University

Lynn E. O'Connor
The Wright Institute

Nathaniel G. Wade
Virginia Commonwealth University

Forgivingness is the disposition to forgive interpersonal transgressions over time and across situations. There is currently no acceptable measure of forgivingness for use in testing theoretical propositions. The authors describe a five-item scenario-based scale, the Transgression Narrative Test of Forgivingness (TNTF). In five studies examining 518 university students from three disparate universities, the authors assess the item and full-scale functioning of the TNTF and its concurrent and 8-week predictive validity relative to trait anger, rumination, neuroticism, agreeableness, and hostility. Test-retest reliability and stability of item locations were both good. Norms are presented by gender, ethnicity, and religious activity. The TNTF is a brief measure of forgivingness that is not theory dependent and is therefore useful in basic and intervention research from a variety of theoretical perspectives.

Over the past three decades, research on forgiveness has been sporadic (see McCullough, Exline, & Baumeister, 1998, for an annotated bibliography). Only in the last few years has the study of forgiveness been characterized by sustained interest among researchers, communication across disciplines, theoretical debate, and an emphasis on theory-driven empirical research (Enright & North, 1998; McCullough, Pargament, & Thoresen, 2000; Worthington, 1998a).

Most previous research has sought to understand the processes of change leading to forgiveness, in which for-giveness has been treated as a state-like dependent variable (for a review and model, see Worthington & Wade, 1999). Research on the processes of forgiveness includes experimental studies in social psychology (see McCullough, Exline, & Baumeister, 1998, for an annotated bibliography) and intervention studies aimed at promoting an act of forgiveness (for reviews, see Hargrave & Sells, 1997; Thoresen, Luskin, & Harris, 1998; Worthington, Sandage, & Berry, 2000). In both the social and applied research traditions, the focus has been on forgiving a single transgression or specific person.

NEEDED—GENERALIZATION ACROSS SITUATIONS AND TIME

Because forgiveness can potentially alleviate emotional distress, it is understandable that researchers have focused on acts of forgiveness. But because of this focus, differences in people's disposition to forgive have gone largely unstudied. Recently, researchers have stressed

the need for studying forgiveness at the dispositional level (Emmons, 2000; McCullough, 2000; Sandage, Worthington, Hight, & Berry, 2000; Worthington, 1998b). We refer to the disposition to forgive as "forgivingness," a term suggested by Roberts (1995) to distinguish the personal trait from acts of forgiving. Forgivingness is a tendency to forgive transgressions that is stable over time and across situations.

Much discussion of forgiveness, its desirability, and its consequences is, at least implicitly, about a disposition to forgive. When asked, about 50% of the U.S. population says they are forgiving (Gorsuch & Hao, 1993). Most are describing their forgivingness rather than referring to a single transgression. Forgivingness may be associated with long-term benefits to social adjustment and to physical and mental health (Kaplan, 1992; Thoresen, Harris, & Luskin, 2000; Williams, 1989). Researchers also have suggested that dispositional as well as contextual variables must be studied to increase our understanding of common human interpersonal problems (Davila & Bradbury, 1998) and their treatment (DiBlasio, 1998). Forgivingness might be one such dispositional variable. The study of dispositional forgivingness, then, could potentially be useful in promoting a variety of positive personal and social outcomes.

DISPOSITIONAL FORGIVINGNESS AND HYPOTHESIZED RELATED TRAITS

Traits Negatively Associated With Forgivingness

Most researchers suggest that dispositional forgivingness promotes long-term benefits to physical health and psychological well-being by overcoming negative traits and affects. Anger is usually described as the main emotional obstacle to forgiveness, the emotion that must be overcome if forgiveness is to occur (Enright, Gassin, & Wu, 1992). Several authors have suggested that dispositional forgivingness should be negatively associated with trait anger, chronic resentment, and hostility, traits with negative consequences for mental and physical health (Kaplan, 1992; Roberts, 1995; Williams, 1989). Empirical studies have supported the relationship between individual acts of forgiveness and the reduction of anger (Huang & Enright, 2000; Weiner, Graham, Peter, & Zmuidinas, 1991). No studies have yet tested whether trait anger and hostility are associated with dispositional forgivingness.

Researchers also have hypothesized that forgivingness is negatively associated with the Big Five personality factor of Neuroticism (Ashton, Paunonen, Helmes, & Jackson, 1998; McCullough, 2000). Neuroticism describes a general tendency to worry and experience negative affect, such as anxiety, depression, and hostility (John, 1990). Ashton et al. (1998) found that emotional stability (the reverse of neuroticism) was correlated positively with forgiveness/nonretaliation. Both McCullough (2000) and Worthington (1998b) have suggested that forgivingness will be correlated with proneness to vengeful rumination (and therefore with neuroticism). No studies to date have tested the relationship between forgivingness and vengeful rumination.

Traits Positively Associated With Forgivingness

Many researchers have hypothesized that acts of forgiveness are facilitated by positive, prosocial affects such as love, compassion, trust, empathy, or sympathy for the transgressor (McCullough, Worthington, & Rachal, 1997; Worthington & Wade, 1999). Both empathy and relationship closeness have been found to increase forgiveness of specific transgressions (McCullough, Rachal, et al., 1998). Researchers have hypothesized that forgivingness is positively associated with the Big Five factor of Agreeableness (Emmons, 2000; McCullough, Rachal, et al., 1998; Worthington & Wade, 1999), which is related to empathy and traits linked to the maintenance of positive interpersonal relationships (Asendorpf, 1998; Graziano, Jensen-Campbell, & Hair, 1996). John (1990) presents data indicating that the trait adjective *forgiving* is associated with agreeableness. Ashton et al. (1998) found that agreeableness correlated positively with both empathy/attachment and forgiveness/nonretaliation.

NEEDED—EFFECTIVE MEASURES OF CONSTRUCTS RELATED TO FORGIVINGNESS

McCullough, Rachal, and Hoyt (2000) have categorized measures of forgiveness as either (a) offense-specific (i.e., forgiveness of a specific person for a specific transgression), (b) dyadic (i.e., forgiveness of a single person for multiple transgressions), or (c) dispositional (i.e., forgivingness). Almost all existing measures are offense-specific or dyadic (Hargrave & Sells, 1997; McCullough, Rachal, et al., 1998; Subkoviak et al., 1995; Wade, 1989). Only one measure, a subscale of a large, broadband assessment inventory, purportedly measures forgivingness (Mauger et al., 1992).

Researchers differ in how they conceptualize forgiveness, and these differences are reflected in the content domains of items included in existing scales. For example, some measures emphasize motivations (McCullough, Rachal, et al., 1998); both cognition and motivations (Wade, 1989); or cognition, affect, and behavior (Hargrave & Sells, 1997; Subkoviak et al., 1995) thought to underlie forgiveness. Although we encourage the construction of theory-based measures in forgiveness research, we believe it is important to have at least a few "ecumenical" measures that can be used by researchers working from diverse theoretical perspectives. Currently, researchers need an ecumenical mea-

sure. Such a measure would be (a) relatively neutral with respect to theoretical definitions of forgiveness; (b) brief enough for intervention, large-scale national survey, or laboratory researcher; (c) psychometrically acceptable; (d) stable across diverse samples of research participants; (e) unbiased with respect to ethnic or gender subgroups; and (f) validated against trait variables that, on their face, seem associated with unforgivingness or forgivingness.

DEVELOPMENT OF THE TRANSGRESSION NARRATIVE TEST OF FORGIVINGNESS (TNTF)

We take our working definition of forgivingness from Roberts (1995), who defines it simply as "an enduring disposition to the act or process of forgiveness" (p. 289). We assume that a person high in forgivingness is *ceritus paribus*, more likely than a person low in forgivingness to forgive an offender for any transgression. With the TNTF, participants imagine themselves experiencing five transgressions and indicate their likelihood of forgiving each offender. The nature and context of the transgressions are standardized, which controls the exact offense, its consequences, and the respondent's relationship with the offender. Although we specify particular events, we recognize that participants interpret scenarios differently, including their sensitivity to transgressions, anger or fear reactions, attitudes toward forgiveness, and beliefs about the motives of the transgressor. Such responses might influence forgiveness in a respondent's daily life and are the variables that must be assessed independently in research on the nature of forgivingness.

We developed items for the TNTF by drawing from the research literature on interpersonal transgressions (see the appendix). Five transgression scenarios, all of which had been used as independent variables in experimental studies, were modified to suit our research purposes. This allowed a conceptual connection with existing research into transgressions. Items 1 and 3 are roughly based on the "Plagiarism" and "Late Term Paper" scenarios used in Gonzales, Manning, and Haugen (1992). Items 2, 4, and 5 are roughly based on Schonbach's (1990) "Neglected Supervision," "Breach of Trust," and "Dubious Self-Defense" scenarios, respectively. These scenarios, with various adaptations of them, have been used in research on facework and offender accounts of transgressions (e.g., Gonzales, Haugen, & Manning, 1994; Hodgins, Liebeskind, & Schwartz, 1996). In adapting the scenarios, we preserved the central theme of the transgressions. Items were written so that test respondents take the role of the victim. We also modified some circumstantial details of the scenarios; for example, we changed the "Late Term Paper" scenario to a "Late Job Application" scenario to avoid

including two school-related items in the test. To try to create distance between the relative locations of the items, we varied the relationship with the transgressor (acquaintance, friend, relative), the blameworthiness of the transgressor (negligence, intentional act), and whether the transgressor apologized (Boon & Sulsky, 1997). Two items (Items 1 and 4) reflect intentional transgressions by acquaintances, two reflect negligent transgressions by friends (Items 2 and 3), and one reflects an intentional transgression by a relative followed by an apology (Item 5). Participants are instructed to read each scenario and circle the number that reflects their likelihood of forgiving the transgression, from 1 (*definitely not forgive*) to 5 (*definitely forgive*) (see the appendix).

THE MEASUREMENT MODEL

Personality psychologists have recently begun to urge the use of item-response scaling in developing measures of personality (Revelle, 1995). In examining the psychometric viability of the TNTF, we employed Rasch scaling procedures (Fischer & Molenar, 1995; Rasch, 1960). Rasch scaling is used to estimate a person's probable response to a test item, taking into account both (a) the degree to which the person possesses the trait being measured and (b) the location of the test item, from easy to endorse to hard to endorse, on a linear continuum. Rasch scaling grades each test item along a linear continuum and uses this continuum as a "yardstick" with which to measure test respondents on the latent variable of interest. Item locations and measures of each person are in the same unit of measurement. Several Rasch measurement models are available for application to polytomous item-response data (Andersen, 1995). Andrich's model for rating scales seems most appropriate for the test format of the TNTF, which has an identical response key shared by all items (Andrich, 1978).

STUDY 1: INITIAL ITEM SCALING

Introduction

We conducted a pilot study to check the dimensionality of the items of the TNTF, assess the fit to the Rasch model, and estimate the reliability of the TNTF.

Method

PARTICIPANTS AND PROCEDURES

We administered the TNTF to 88 undergraduate students at an urban, mid-Atlantic state university. The students were 68 women (76%) and 20 men (23%), with a mean age of 20.1 years ($SD = 6.2$). The ethnic identities of the students were as follows: 44 European Americans

TABLE 1:   Item and Scale Statistics for the Transgression Narrative Test of Forgivingness (TNTF)

Sample	Initial Scaling (Study 1)				Validation (Study 2)				Cross-Validation (Study 3)				Combined			
Item	d (SE)	fit_W	fit_U	r_c	d (SE)	fit_W	fit_U	r_c	d (SE)	fit_W	fit_U	r_c	d (SE)	fit_W	fit_U	r_c
1	1.03 (.17)	1.37	1.30	.32	.48 (.10)	.92	.91	.54	.79 (.09)	1.06	1.05	.60	.72 (.06)	1.07	1.05	.55
2	−.46 (.14)	1.15	1.14	.47	−.40 (.12)	1.34	1.30	.43	−.75 (.10)	1.32	1.28	.50	−.57 (.07)	1.29	1.26	.48
3	−.31 (.14)	.66	.66	.59	−.12 (.10)	.76	.77	.56	−.17 (.09)	.93	.92	.60	−.18 (.06)	.82	.82	.60
4	.39 (.14)	.88	.95	.53	.55 (.10)	.95	.93	.54	.67 (.09)	.82	.82	.64	.57 (.06)	.88	.89	.59
5	−.65 (.14)	.95	.93	.53	−.50 (.11)	1.07	1.03	.55	−.54 (.09)	.90	.89	.61	−.54 (.06)	.96	.94	.58

NOTE: d is the estimated item location; SE is the standard error of the item location; fit_W is the weighted mean-square-fit statistic; fit_U is the unweighted mean-square-fit statistic; and r_c is the corrected item-total correlation. Total scale statistics are as follows: for "Initial Scaling," $n = 88$, alpha $= .73$, Rasch Person $R = .78$, Rasch Item $R = .95$; for "Validation," $n = 146$, alpha $= .76$, Rasch Person $R = .81$, Rasch Item $R = .95$; for "Cross-Validation," $n = 233$, alpha $= .81$, Rasch Person $R = .83$, Rasch Item $R = .98$; and for "Combined Samples," $n = 467$, alpha $= .79$, Rasch Person $R = .82$, Rasch Item $R = .99$.

(49%), 29 African Americans (33%), 6 Asian Americans (7%), 4 Hispanic Americans (5%), and 6 others (7%). Students participated voluntarily for class credit. We announced the study in classes. Participants picked up a questionnaire at a designated location, completed it, and placed it in a return box. Of the 100 picked up, 88 usable questionnaires were returned.

Results

On the TNTF, raw score totals ranged from 5 to 25. The mean for the sample was 13.3. The standard deviation was 3.5. To screen for the dimensionality of the items, we extracted the eigenvalues of the correlation matrix of item raw scores using a principal components analysis with varimax rotation. Smith (1996) has advised that if raw test scores are dominated by a single eigenvalue, then Rasch scaling of items can be performed efficiently. The first eigenvalue (2.41) clearly dominated the second largest eigenvalue (.90).

The items were then fit to the Rasch rating scale model using joint (unconditional) maximum likelihood estimation. In the note to Table 1 we provide estimates of three summary reliability statistics for the TNTF: (a) Cronbach's alpha ( ), (b) Rasch asymptotic person separation reliabilities (Person R), and (c) Rasch asymptotic item separation reliabilities (Item R). The Person R is an upper limit of the proportion of trait variance not attributable to measurement error. It is analogous to Cronbach's alpha and is interpreted accordingly. The Person R was .78; alpha was .73. Both of these statistics suggest an acceptable amount of variability among participants on forgivingness. The Item R is an estimate of the proportion of item variance that is free of estimation error. It reflects both the spread of item locations and their standard errors. An acceptable Item R ($> .90$) indicates that test items are well separated, with sufficiently small estimation errors. Thus, they are a useful "yardstick" for measuring the variable of interest. The Item R

was .95, suggesting an acceptable amount of variability in item locations.

We present statistics for each item of the TNTF in Table 1. First, we present item location estimates, d, which are in logit units. Higher values indicate more difficult-to-endorse items. The mean of the item locations is conventionally set to zero. Item 5 (dubious self-defense) had the lowest item location ($d = -.65$); Item 1 (plagiarism) had the highest location ($d = 1.03$). Thus, participants were most likely to forgive in Item 5 and least likely to forgive in Item 1. Second, we present approximate standard errors, s, of the location estimates. Third, we present weighted mean-square-fit statistics, fit_W, which provide a measure of fit for items nearest the center of item locations. This fit statistic has an expected value of 1.00. Wright and Linacre (1994) suggest that values between .60 and 1.40 are reasonable for rating scales; all were within the acceptable range (.66 to 1.37). Fourth, we present unweighted mean-square-fit statistics, fit_U, which are sensitive to outlying responses. These statistics have an expected value of 1.00 and are evaluated by the same criteria as fit_W; all were within the acceptable range (.66 to 1.30). Fifth, we present the corrected item-total correlations; that is, Pearson's product-moment correlations between each item and the sum of the other items.

Discussion

The Cronbach's alpha and Rasch reliabilities for the TNTF from the initial scaling were all acceptably high, especially for a five-item scale and considering the relatively small sample size used for the estimation of model parameters. These results indicate that both the item locations and person trait measures had sufficient variability and were estimated with a tolerable amount of measurement error. The fit statistics for individual items suggest that the items of the TNTF are adequately unidimensional and are ordered sufficiently in locations. Given these results, we retained all of the initial

items and used them in the validation studies described below.

STUDY 2: VALIDATION STUDY

Introduction

In this study, we first replicated the Rasch modeling in a new sample. We also tested the concurrent construct validity of the TNTF in this new sample. Researchers and theorists have suggested that forgivingness should be negatively associated with chronic anger and related variables such as resentment and hostility (Kaplan, 1992; Williams, 1989). Emmons (2000), Worthington (1998b), and McCullough (2000) have proposed that neuroticism is negatively and agreeableness positively associated with forgivingness. In the present study, we expect that to provide evidence for the convergent construct validity of the TNTF, we should find moderate negative correlations of the TNTF with measures of trait anger, hostility, and neuroticism and a moderate positive correlation with agreeableness. Evidence for discriminant construct validity would be indicated by nonsignificant correlations with conscientiousness, extraversion, and openness to experience. We also do not anticipate correlations with the expressive components of aggression (physical and verbal aggression).

Method

PARTICIPANTS

Participants were undergraduates ($N = 146$) from a state university in the San Francisco area ($n = 66$) and a private religious university in the Pacific Northwest ($n = 80$). Data were collapsed into a single data set to allow a sample large enough to support statistical analyses. The sample included 111 women (76%) and 35 men (24%). Participants' mean age was 19.5 years ($SD = 2.75$). The ethnic composition of the sample was as follows: 75 European Americans (51%), 2 African Americans (1%), 38 Asian Americans (26%), 16 Hispanic Americans (11%), and 13 others (9%). Students at both universities participated voluntarily for class credits.

INSTRUMENTS

Demographic Questionnaire (DQ). Participants reported their gender, ethnicity, age, and income.

The Trait Anger Scale (TAS) (Spielberger, Jacobs, Russell, & Crane, 1983). The TAS is a 15-item scale for assessing anger as a personality trait, which is defined in terms of the frequency of angry states over time. Alpha coefficients among college students were .87 for men and .87 for women. Among Navy recruits, alpha coefficients were .87 for men and .84 for women. In addition to a total score, the TAS can be scored for Angry Temperament and Angry Reaction subscales, each consisting of

four items. The Angry Temperament subscale assesses anger without specifying any particular anger-provoking circumstance. The Angry Reaction subscale refers to specific frustrations or negative evaluations.

The Aggression Questionnaire (AQ) (Buss & Perry, 1992). The AQ is a 29-item self-report questionnaire used to assess four dimensions of aggression: anger, hostility, verbal aggression, and physical aggression. The validity of the AQ has been supported by correlations with competition, assertiveness, impulsiveness, and peer ratings of aggressiveness. The internal consistency of the AQ has been estimated to be .89 (total scale) and from .72 to .85 for subscales. Buss and Perry (1992) report 9-week test-retest reliabilities between .72 to .80 for subscales and .80 for the total scale.

The Big Five Personality Inventory, V44 (BFI-44) (John, Donahue, & Kentle, 1992). The BFI-44 is a 44-item measure of the Big Five personality traits: neuroticism, extraversion, agreeableness, conscientiousness, and openness to experience (John, 1990). The BFI-44 consists of short phrases that are rated on a Likert-type scale according to how descriptive the phrases are of the respondent. Each subscale consists of 8 to 10 items. John et al. (1992) report internal consistencies for the subscales ranging from .75 to .88. The subscales were further validated by findings of moderate correlations between peer-peer reports and peer-self reports, which ranged from .21 for agreeableness to .63 for extraversion.

PROCEDURE

Packets of questionnaires were distributed in classes. Students returned completed packets at the following class. The return rate was 66% at the university in San Francisco and 80% at the university in the Pacific Northwest. Students in San Francisco completed only the TNTF and TAS. Students in the Pacific Northwest completed the TNTF, TAS, AQ, and BFI-44.

Results

REPLICATION OF ITEM AND SCALE CHARACTERISTICS

The items of the TNTF were fit to the Rasch rating scale model. The first eigenvalue (2.56) dominated the second largest eigenvalue (.78). The alpha (.76), Rasch Person R (.81), and Rasch Item R are comparable in magnitude to those obtained in Study 1. We report the item characteristics in Table 1. Items 4 (breach of trust) and 1 (plagiarism) reversed location order from the initial sample, indicating that participants were least likely to forgive in Item 4 in this sample, whereas participants were least likely to forgive Item 1 in the initial sample. The remaining item locations mirrored results from Study 1. The weighted mean-square-fit statistics ranged

TABLE 2:   Correlations Between the Transgression Narrative Test of Forgivingness (TNTF) and Validation Scales

Trait Anger Scale	Validation (Study 2)			Cross-Validation (Study 3)			Prospective (Study 4)	
	r	n	M (SD)	r	n	M (SD)	r	n
Temperament[a]	−.37***	142	6.4 (2.6)	−.33***	233	6.4 (2.5)	−.44***	60
Reactivity[a]	−.36***	143	9.3 (2.8)	−.34***	232	9.1 (2.7)	−.27*	60
TAS total[a]	−.43***	141	29.5 (7.9)	−.38***	230	28.9 (7.5)	−.43***	60
Aggression Questionnaire								
Anger[a]	−.33**	78	13.9 (4.4)	−.35**	56	13.5 (3.4)	−.45***	61
Hostility[a]	−.21	78	18.9 (6.0)	−.37**	56	17.1 (4.7)	−.32*	61
Verbal aggression	−.13	76	13.1 (3.7)	−.15	56	12.1 (3.0)	−.12	60
Physical aggression	−.09	78	15.2 (6.0)	−.21	55	14.9 (4.8)	−.24	61
Big Five Inventory								
Agreeableness[a]	.25*	80	3.9 (.56)	.33***	232	3.9 (.63)	.28*	61
Neuroticism[a]	−.29*	80	2.8 (.73)	−.27***	232	2.9 (.77)	−.32*	61
Conscientiousness	.15	80	3.9 (.59)	.24***	232	3.8 (.63)	.16	61
Extraversion	−.02	80	3.4 (.93)	.19**	229	3.4 (.81)	−.02	61
Openness	.14	80	3.8 (.62)	.14	230	3.7 (.66)	.02	61
Dissipation-Rumination Scale (DRS)[a]							−.49***	61
Social desirability				−.06	55	51.9 (4.7)		

NOTE: A modified Bonferoni correction of $p = .01$ was accepted as statistically significant. Means and standard deviations for Study 4 validation scales, except for the DRS ($M = 22.9$, $SD = 10.7$), are the same as in Study 3; only the DRS and TNTF were administered at an 8-week retest. Means and standard deviations for the TNTF are as follows: for Study 2, $M = 15.5$, $SD = 3.9$; for Study 3, $M = 15.5$, $SD = 4.0$; and for Study 4, $M = 16.6$, $SD = 3.3$.
a. Theory-relevant hypothesized correlations.
*$p < .05$, ns. **$p < .01$. ***$p < .001$. .

from .76 to 1.34, and the unweighted mean squares ranged from .77 to 1.30.

EVIDENCE OF VALIDITY

The construct validity of the TNTF was tested by computing Pearson's product-moment correlations between the validation scales and the Rasch logit measures on the TNTF. These correlations are displayed in Table 2, under the heading "Validation." Two-tailed significance tests were used. Because we tested seven theory-relevant hypotheses in this study, we used a modified Bonferroni correction of $p < .01$ to protect against inflated Type I error. As hypothesized, the TNTF correlated significantly in the predicted directions with all subscales of the TAS and with the Anger subscale of the AQ. The TNTF was not significantly correlated ($p < .01$) with the Hostility, Verbal Aggression, or Physical Aggression subscales of the AQ, nor was the TNTF correlated significantly with any of the Big Five personality traits.

Discussion

The psychometric stability of the TNTF was supported. The scale and item statistics were similar to those obtained in the initial scaling sample. We obtained a slightly larger Person R and Cronbach's alpha in the present sample, perhaps due to the larger and more ethnically diverse sample used for Study 2. Item statistics in Table 1 were similar to those obtained previously, but the overall fit of items was better in the present sample.

Pearson's correlation between item location estimates in the initial scaling sample and the validation sample was .93. This suggests that the relative item locations remained stable in two samples that are geographically, ethnically, and perhaps culturally different.

Evidence for convergent validity was indicated by strong relationships between the TNTF and the TAS total score, Angry Temperament, Angry Reaction, and the Anger subscale of the AQ. A person who is dispositionally inclined toward anger is likely to be inclined toward unforgivingness (Williams, 1989). Contrary to hypothesis, we found no significant relationship between forgivingness and hostility (on the AQ). The weak correlation (−.21) casts some doubt on the assertions of some health psychologists that low trait forgivingness might be related to cardiovascular problems in middle age or later life (Kaplan, 1992; Thoresen & Goldberg, 1998; Williams, 1989). Free-floating hostility seems to be related to poor cardiovascular health. If forgivingness is not related to hostility in replications of the present study, then the association between forgivingness and health would be dubious.

Contrary to expectations, forgivingness was not significantly correlated with agreeableness or neuroticism. Forgivingness was correlated .25 with agreeableness and −.29 with neuroticism. These correlations are similar in magnitude to those reported by Ashton et al. (1998) between forgiveness/ nonretaliation and agreeableness (.29) and emotional stability (−.21), which is the oppo-

site pole of neuroticism. However, because of the low number of participants who completed the AQ in the present study, the correlations were not statistically significant. The magnitude of the correlations for forgivingness with both agreeableness and neuroticism makes additional investigation of the relationship necessary. Evidence for discriminant validity was indicated by nonsignificant correlations between the TNTF and conscientiousness, extraversion, openness, and physical and verbal aggression, as anticipated.

STUDY 3: CROSS-VALIDATION

Introduction

We replicated Study 2 using a larger sample, providing a cross-validation of the structure and construct validity of the TNTF. We used three new data collection efforts, which we combined into a single sample. Although the correlation between conscientiousness and forgivingness in the previous study was small, we were concerned, based on McCullough, Rachal, et al. (1998), that a constellation of positive behaviors and even a social desirability response set might influence the relationship between forgivingness and other variables; therefore, we thought it worthwhile to include a measure of social desirability in a religious subsample. We also included the AQ with that subsample.

Method

PARTICIPANTS

Participants for the cross-validation study were undergraduates (N = 233) from the same mid-Atlantic state university used in Study 1 (n = 177) and the same Pacific Northwest, private, Christian university used in Study 2 (n = 56). Samples from the different universities were combined to provide a large enough sample to permit adequate statistical analyses. There were 193 female participants (83%) and 40 male participants (17%), with a mean age of 23.4 (SD = 7.6). The sample included 144 European Americans (62%), 54 African Americans (23%), 17 Asian Americans (7%), 8 Hispanic Americans (3%), and 9 participants with other ethnic identities (4%). All participated voluntarily for class credit.

INSTRUMENTS

The Marlowe-Crowne Social Desirability Scale (MCSD) (Crowne & Marlowe, 1960). The MCSD is a well-known, 33-item, self-report scale originally constructed to measure the tendency to respond in a socially desirable manner to psychological tests. Other research has linked the MCSD to such motivational and impression management constructs as need for approval and the avoidance of disapproval (Crowne, 1979). The internal consistency

of the items and test-retest reliability have been established in many studies.

The DQ, TAS, and BFI-44 all described in Study 2 were administered to all participants. The 56 participants from the Christian university also completed the AQ and MCSD.

PROCEDURE

Packets of questionnaires were distributed in classes, and participants returned their completed questionnaires at the following class meeting. The response rate was 88% at the mid-Atlantic university and 93% at the Pacific Northwest university.

Results

SECOND REPLICATION OF ITEM AND SCALE CHARACTERISTICS

The items of the TNTF were fit to the Rasch model. Statistics for the scale are displayed in the note to Table 1. The first eigenvalue (2.80) dominated the second (.70). The Cronbach's alpha and Rasch Person R were larger than in the two previous samples, which is to be expected given the larger sample size. The Item R of .98 is large for a five-item scale. The item statistics are shown in Table 1. Again, all items fit the Rasch model adequately. Item locations paralleled those estimated in Study 1 for Item 1 (plagiarism) and Item 4 (breach of trust). Items 2 (neglected supervision) and 5 (dubious self-defense) were reversed in order relative to Studies 1 and 2. To assess the stability of item locations, we computed Pearson's correlations between the calibrations estimated in Study 2 and Study 3 (r = .98) and Study 1 and Study 3 (r = .94). The item estimates thus appear to be stable across diverse samples.

CROSS-VALIDATION: MORE EVIDENCE OF VALIDITY

The results of the cross-validation are shown in Table 2. As in Study 2, Pearson's correlations were calculated between the validation scales and the Rasch logit measures of the TNTF. Again, two-tailed significance tests were used with a testwise alpha level of .01. Correlations were statistically significant in the hypothesized directions for all subscales of the TAS and with the Anger and Hostility subscales of the AQ. The TNTF was significantly correlated with agreeableness and neuroticism. The TNTF was not significantly correlated with extraversion and openness to experience, showing discriminant validity. Contrary to hypothesis, the TNTF was correlated with conscientiousness. The correlation between the TNTF and the MCSD for participants at the Christian university was virtually zero.

Discussion

In Study 3, we replicated the structure of the TNTF. The scale reliabilities, item structure, convergent validity, and discriminant validity were largely in line with theoretical expectations. In both Study 2 and Study 3, the TNTF correlated negatively with the TAS total score and the Angry Temperament and Angry Reaction subscales. It also correlated negatively with the Anger subscale of the AQ in both studies. The correlations are comparable in magnitude across both studies. These findings are consistent with theorizing about the nature of forgivingness (Kaplan, 1992; Roberts, 1995; Williams, 1989). These correlations represent the minimum necessary evidence of construct validity of the TNTF.

In contrast with Study 2, the TNTF also correlated with the Hostility subscale of the AQ. This finding is consistent with predictions based on Kaplan (1992) and Williams (1989). However, the difference in the magnitudes of the correlations in Study 2 (−.21) and Study 3 (−.37) is substantial. The cumulative evidence for a relationship between forgivingness and hostility remains inconclusive. The Hostility subscale of the AQ contains items that reflect a variety of negative attitudes toward others ("I am sometimes eaten up with jealousy," "At times I feel I have gotten a raw deal out of life," "Other people always seem to get the breaks"). Not all generalized negative attitudes toward others are likely to be equally related to forgivingness. Most of the hostility items are not specifically tied to interpersonal transgressions but are more complex attitudes that may not be associated strongly with forgiveness of specific offenses. Therefore, although general hostility is related to health consequences, it is less clear that hostility is consistently (negatively) associated with forgivingness. Further evidence is needed before we can confidently suggest a negative relationship between forgivingness and hostility.

The correlations between the TNTF and both agreeableness and neuroticism were similar in the validation and cross-validation samples. Given the larger sample size of the cross-validation study, these correlations were significant, as predicted by Worthington (1998b) and McCullough (2000).

STUDY 4: TEST-RETEST STABILITY, PREDICTIVE
VALIDITY, AND FURTHER CONVERGENT VALIDITY

Introduction

Studies 2 and 3 provided evidence for the concurrent validity of the TNTF. In trait measures, though, concurrent validity is necessary but not sufficient evidence for the validity of an instrument. We also need to establish the TNTF's temporal consistency and ability to predict other trait measures that are assessed in some distal mea-

surement. We readministered the TNTF after 8 weeks to one of the subsamples reported in Study 2.

Forgivingness has been hypothesized to be related to a tendency to ruminate about transgressions (McCullough, 2000; McCullough, Rachal, et al., 1998; Worthington, 1998b; Worthington & Wade, 1999). Rumination about transgressions has been linked to a variety of cognitive, emotional, and behavioral variables associated with aggression (Caprara, Manzi, & Perugini, 1992; Collins & Bell, 1997). Rumination is especially likely when a person experiences negative emotions. A tendency to ruminate about negative events can perpetuate negative affect associated with those events (Greenberg, 1995). Following an aversive event, attributions are frequently made about causes of negative emotions (Weiner, 1985). In our view, vengeful rumination that attributes the causes of a transgression to stable characteristics of a transgressor is likely to be associated with unforgiveness. We administered the Dissipation-Rumination Scale (DRS) (Caprara, 1986) at the 8-week posttest. This permitted a concurrent correlation (at Time 2) of the TNTF and DRS and a predictive correlation (using the TNTF at Time 1 to predict the DRS score at Time 2).

Finally, we asked partners in romantic relationships to rate themselves and their partner on forgivingness. Correlation of self-other ratings is another criterion against which validity can be assessed.

Method

PARTICIPANTS

Of the 146 participants in Study 2, we examined 62 of the 80 participants (78%) in the subsample from the Pacific Northwest university. Of the 62, 50 were women (81%) and 12 were men. There were 53 European Americans (84%), 4 Asian Americans (6%), 2 Hispanic Americans (3%), and 4 others (7%).

In addition, we surveyed 26 couples who were married (*n* = 8 partners) or dating (at least 3 dates; *n* = 45 partners). One rating was incomplete, yielding 51 partners making ratings. The mean age of partners was 24.4 (*SD* = 5.5). Of the 51 partners, 20 were European American, 25 were African American, 3 were Hispanic American, 1 was Asian American, and 2 were other ethnicities.

INSTRUMENTS

The TNTF, TAS, AQ, and BFI-44 were administered in Study 2 (Time 1). The DRS was administered at Time 2, 8 weeks after the first administration, and the TNTF was readministered.

The DRS (Caprara, 1986). The DRS is a 20-item, self-report scale used to assess the effect over time of an individual's desire to act aggressively following an interpersonal offense. Of the 20 Likert-type items, 5 are "control"

items that are not scored. Dissipation and rumination are conceptualized as opposite ends of a continuum. Dissipation reflects the tendency to "shrug off" insults or offenses with little rumination. Rumination reflects the tendency toward an increasing desire for retaliation over time. The internal consistency of the English version of the scale was estimated to be .87. The scale has been validated using structural models and correlations with such variables as tolerance toward violence, irritability, emotional susceptibility, and fear of punishment (Caprara et al., 1992). Validity has been supported in studies in which participants were given the opportunity to retaliate against research confederates who insulted them (Caprara, 1986).

PROCEDURE

Procedures for administration of the TNTF, TAS, AQ, and BFI-44 were described in Study 2. Eight weeks after that administration, all 80 participants were given the opportunity to complete the TNTF (again), the DRS, and several questionnaires associated with a different investigation. For the validation using partner ratings, couples came to a testing location, were separated, and completed the TNTF concerning both themselves and their partner.

Results

EIGHT-WEEK TEST-RETEST RELIABILITIES

The measures of the TNTF from the two measurement occasions had an 8-week test-retest reliability of $r(60) = .69$, $p < .001$. The correlation between the item locations obtained on the two testing occasions was .99. Person measures and item locations were stable over a 2-month testing interval.

EIGHT-WEEK PREDICTIVE VALIDATION AND FURTHER CONCURRENT VALIDATION

To test the 8-week predictive validity of the TNTF, we calculated Pearson's correlations between the TNTF measures obtained from the second testing with the validation scales that had been completed 2 months earlier. The DRS, which was administered at the follow-up testing, was correlated with TNTF measures obtained 2 months earlier. Two-tailed significance tests were used and the testwise alpha was .01. These correlations are shown in Table 2 under the heading "8-Week Predictive Validation." Because both the TNTF and the DRS were administered together at the second testing, their correlation was a test of concurrent validity, $r(60) = -.46$, $p < .001$. In addition, the TNTF at Time 1 was used to predict the DRS 8 weeks later, $r(59) = -.49$, $p < .001$. Thus, the correlations between the TNTF and rumination were statistically significant both predictively and concurrently.

The results of the predictive validation are similar in magnitude to those in both the validation and cross-validation studies (see Table 2). The TNTF was significantly correlated with total TAS and Anger Temperament and the Anger subscale of the AQ. The TNTF was not significantly correlated with Anger Reaction (TAS) or with hostility or physical and verbal aggression (AQ). No significant relationships were found between the TNTF and BFI-44 scales.

VALIDATION AGAINST RATING BY A RELATIONSHIP PARTNER

The Pearson's correlation between ratings of self and by partners was $r(49) = .60$, $p < .001$.

Discussion

Study 4 provides further support for the stability and utility of the TNTF as a psychometric instrument. The 8-week test-retest reliability of forgivingness measures was moderately high. Ratings of self were similar to ratings by a relationship partner. The results of Study 4 support the construct of forgivingness as a stable personality disposition related meaningfully to measures of anger. For hostility, agreeableness, and neuroticism, we found correlations with the TNTF that, although not statistically significant at the conservative .01 alpha level, were comparable in magnitude to those obtained in the concurrent validity studies. Across Studies 2, 3, and 4, there is weak support for a relationship between forgivingness and hostility, agreeableness, and neuroticism. However, rumination was significantly correlated with the TNTF, concurrently and predictively. This result is consistent with predictions of Worthington (1998b) and McCullough (2000) and with previous research using a different measure of rumination that predicted forgiveness of single transgressions (McCullough, Rachal, et al., 1998).

STUDY 5: DIFFERENTIAL ITEM FUNCTIONING AND NORMATIVE DATA

Introduction

Ideally, the items of a test should have the same locations in any subgroup of test respondents, such as gender or ethnic subgroups. The yardstick should remain the same, regardless of the average trait level of the subgroups being measured. If the measuring instrument varies between subgroups, then the subgroups cannot be reliably compared.

To test for differential item functioning using Rasch methods, items are calibrated separately for all subgroups. The items are then equated onto the same scale and item locations are compared between subgroups (Draba, 1977). For example, item locations for women (d_F) can be subtracted from item locations for men (d_M),

yielding a difference $(d_M\text{-}d_F)$. A positive difference would indicate that women found the item more difficult to endorse than did men.

In this study, we examine gender and ethnic subgroups for differential item functioning. If differential item functioning is not indicated, then we may compare subgroup differences in forgivingness without undue concern about potentially biased test items. We therefore provide normative data (and statistical comparisons of subgroup differences) for gender and ethnic subgroups after the differential item functioning analyses. Based on a meta-analysis of forgiveness intervention research, we anticipated no difference in forgivingness between men and women (Worthington et al., 2000). For ethnicity, we were uncertain about whether to expect differences. One reason for our uncertainty is that ethnicity and religion are often intertwined. African Americans have been found to be more religious than European Americans (Levin & Taylor, 1998), and religion has been related to being a forgiving person, both theoretically (McCullough & Worthington, 2000) and empirically (Gorsuch & Hao, 1993). However, forgiveness is not confined to religious people.

We examined the relationship between forgivingness and other demographic variables. Enright and his colleagues (1992) have shown that reasoning about forgiveness develops from childhood to early adulthood, but after college age, no age differences in reasoning about forgiveness have been found. Therefore, for our sample of college students, we anticipate no difference according to age. We have found no suggestion in the literature that socioeconomic status should be expected to make a difference in forgivingness. McCullough and Worthington (2000) have hypothesized that highly religious people are likely to be more forgiving as a subgroup than are less religious people. Participants were asked to indicate their frequency of religious activity. We therefore compared participants who were highly active in their religious communities to those who were less active.

Method

In Studies 1 through 3, we reported five data collection efforts—two from a large, urban, mid-Atlantic public university (n = 265); two from a small, Pacific Northwest, private, Christian university (n = 136); and one from a large, San Francisco, public university (n = 66). The total normative sample included 467 university students. All participants completed a demographic form that included gender, ethnicity, age, monthly income, and religious affiliation. All participants except those from the San Francisco public university were asked, "How many activities or services do you attend at your religious institution?" Response options were "none,"

TABLE 3: Differential Item Functioning of Transgression Narrative Test of Forgivingness (TNTF) Items in Combined Sample by Gender and Ethnicity

	Gender		Ethnicity	
Item	$d_M\text{-}d_F$	$d_{EU}\text{-}d_{AF}$	$d_{EU}\text{-}d_{AS}$	$d_{AF}\text{-}d_{AS}$
1	−.49	.49	.33	−.17
2	.45	−.22	−.43	−.21
3	.26	−.16	.09	.25
4	−.27	.11	−.11	−.22
5	.06	−.23	.12	.35
r	.92	.96	.92	.90

NOTE: M = male, F = female, EU = European American, AF = African American, and AS = Asian American. r = correlation between item locations estimated in separate subgroups.

"one a year," "a few times a year," "one a month," "one a week," and "more than one a week."

Results

DIFFERENTIAL ITEM FUNCTIONING

The differential item functioning analyses by gender and ethnic subgroups are shown in Table 3. We examined only the three ethnic groups that had at least 50 participants in the study: European Americans (n = 262), African Americans (n = 85), and Asian Americans (n = 61). Wright (1996) has suggested that at least 50 test respondents are needed in a single calibration sample to obtain practically useful estimates of item locations. We present the differences in item location estimates obtained from the subgroups. Differences of less than half a logit suggest no serious item bias for most testing situations (Wright & Douglas, 1975). We also present Pearson's correlations between the item location estimates obtained from the separate calibrations.

DEMOGRAPHIC COMPARISONS AND NORMATIVE DATA

The mean TNTF raw score for the combined samples (N = 467) was 14.6 (SD = 3.9). Ages of participants ranged from 18 to 57. Monthly income ranged from $0 to $7,250. The correlation of TNTF Rasch measures with age was $r(465)$ = .13, p < .01, and with income was $r(410)$ = .03, ns.

We calculated a two-way (2×5) analysis of variance to compare gender and ethnic groups on the TNTF. The main effect for gender indicated no statistically significant difference between men (n = 92, M = 14.7, SD = 3.8) and women (n = 371, M = 14.5, SD = 4.1), $F(1, 449)$ = .29, ns. The main effect for ethnic groups was statistically significant, $F(4, 449)$ = 2.84, p < .05. Post hoc analyses using Tukey's Honestly Significant Difference (HSD) tests were used to compare ethnic groups. The European American group (n = 262, M = 15.2, SD = 4.1) was signifi-

cantly higher on forgivingness than were both the Asian American (n = 61, M = 13.4, SD = 3.4) and Hispanic American (n = 28, M = 12.7, SD = 4.2) groups. The African American (n = 85, M = 14.2, SD = 3.8) and Other (n = 27, M = 14.1, SD = 3.8) groups did not differ significantly from each other or any of the other ethnic groups. The interaction between gender and ethnicity was not statistically significant, $F(4, 449)$ = .59, ns.

Because some cell sizes were too small to include religious activity in a factorial analysis with gender and ethnicity, highly active and less active participants were compared on the TNTF using a one-factor ANOVA. Religiously active participants (n = 179, M = 16.5, SD = 3.8) were significantly higher in forgivingness than were less active participants (n = 215, M = 13.4, SD = 3.4), $F(1, 393)$ = 71.46, $p < .001$. We thought it possible that the ethnic group differences reported above were due to differences in religious activity; however, the ethnic groups did not differ significantly in the proportion of high and low religious participation, $\chi^2(4)$ = 3.7, p = .44, nor did the gender groups, $\chi^2(1)$ = .13, p = .80.

We also computed all ANOVAs using Rasch logit measures. The substantive results of the comparisons were the same as the ANOVAs comparing raw scores. (For data on the logit comparisons, and for a conversion table for transforming raw scores to Rasch logits, contact the first author.)

Discussion

There appears to be little evidence for differential item functioning by gender or ethnicity. In the demographic comparisons, the lack of a significant difference in forgivingness between men and women is consistent with findings of no gender effects in response to interventions designed to promote forgiveness (Worthington et al., 2000). Although Worthington et al. found that more female than male students volunteered for studies, this might be because women comprised about 70% to 80% of the students in classes in psychology in participating schools. The small correlations between the TNTF and age and income also were expected. (The correlation with age was statistically significant but trivial in magnitude.) However, the finding that European Americans were higher in forgivingness than both the Asian Americans and Hispanic Americans was not anticipated.

The present study demonstrates that participation in religious services and activities, although a rather unsophisticated measure of religion (see Larson, Swyers, & McCullough, 1998), can nonetheless distinguish between people high and low in forgivingness. The difference between the religiously active participants (those participating in religious activities once a week or more) and the less religiously active participants (those

participating in religious activities once a month or less) is consistent with Gorsuch and Hao (1993).

GENERAL DISCUSSION

We have provided empirical support for the construct of forgivingness, the disposition to forgive interpersonal transgressions over time and across situations. Such a construct has been suggested by many researchers and theoreticians (Mauger et al., 1992; McCullough, 2000; Roberts, 1995; Worthington, 1998b), but little empirical evidence has been adduced (cf. Mauger et al., 1992). Moreover, this disposition seems to be adequately measured by a five-item, narrative-based test, the TNTF. Although brief, the TNTF appears to be reasonably well behaved psychometrically.

Using the TNTF, we have obtained initial evidence that the disposition to forgive transgressions is negatively associated with a variety of negative affective traits. People low in forgivingness might be described as prone to anger, anxiety, and other negative emotions. They tend to ruminate in a vengeful manner following offenses and perhaps hold hostile attitudes toward others. Our data suggest a stable, moderate correlation between forgivingness and agreeableness. This finding is consistent with the view that forgiveness is facilitated by prosocial affects such as empathy, compassion, and trust (McCullough, 2000; Worthington, 1998b).

The present research has at least three major limitations. First, all participants were students enrolled in undergraduate psychology courses. Despite finding stability for the TNTF across five samples from three substantially different universities, we remain tentative in endorsing the TNTF for use with noncollege populations. Second, we used relatively small samples. Rather than combining samples, we considered it more important to demonstrate the consistency of results across diverse samples taken at different measurement occasions. Third, there is a bias inherent in correlating self-report measures. Although we validated the TNTF against ratings by significant others, future research should validate it against behavioral observations of forgiving behavior after a transgression and perhaps physiological correlates of people identified to be high in forgivingness. In addition, the TNTF also should be investigated as a predictor of response to forgiveness-inducing interventions or perhaps as a moderator distinguishing between differential responding to different types of interventions. Despite our failure to consider all of the above different types of validity data within this article, the self-reports (and other-ratings) presented in the present article are sufficiently strong to warrant future studies using the TNTF.

We regard the development of the TNTF as a first step in research on dispositional forgivingness. Because no

adequate measure of forgivingness has been published until now, the field of research on forgivingness is essentially wide open to new research ideas. We are especially hopeful for the cross-fertilization of basic and applied research on forgivingness.

Research on the causes, correlates, and consequences of forgivingness should appeal to a broad community of researchers. Potential areas of research include cognitive, affective, motivational, personality, biological, and genetic factors associated with forgivingness (see Worthington & Wade, 1999, for a review). Cognitive variables that might be associated with forgivingness include implicit theories of others (Dweck, Hong, & Chiu, 1994), typical defenses (Baumeister & Cairns, 1992), processing styles (Caprara, 1986), and generalized attitudes (such as trust or hostility). Many other affective trait variables that might be linked to forgivingness, such as guilt and shame (Tangney, Wagner, Barlow, & Marschall, 1996), warrant research. Possible interpersonal motivations and behaviors associated with forgivingness include personal strivings (McAdams, 1994), attachment styles (Diehl, Elnick, Bourbeau, & Labouvie-Vief, 1998), the need to belong (Baumeister & Leary, 1995), and conflict styles (Van de Viert & Euwema, 1994). A number of developmental variables might affect trait forgivingness, such as moral and religious training, parenting styles, and history of conflict and abuse. The field of research on forgivingness is currently ripe for theoretical, applied, and psychometric study. We hope the TNTF will play a useful role in this developing field.

APPENDIX
Transgression Narrative Test of Forgivingness (TNTF)

Below are a number of situations in which people might find themselves. People respond in different ways to these situations in terms of what things they will forgive. We would like you to read each situation and imagine it has happened to you. Then we would like you to use the scale below to indicate how you think you would respond to the situation:

1 = *definitely not forgive,*
2 = *not likely to forgive,*
3 = *just as likely to forgive as not,*
4 = *likely to forgive,* and
5 = *definitely forgive.*

1. Someone you occasionally see in a class has a paper due at the end of the week. You have already completed the paper for the class and this person says he or she is under a lot of time pressure and asks you to lend him or her your paper for some ideas. You agree, and this person simply retypes the paper and hands it in. The professor recognizes the paper, calls both of you to her office, scolds you, and says you are lucky she doesn't put you both on academic probation. Imagine yourself in such

a situation and mark how likely you are to forgive the person who borrowed your paper.

1   2   3   4   5

2. A fairly close friend tells you that he or she needs some extra money for an upcoming holiday. You know a married couple who needs a babysitter for their 3-year-old for a couple of nights and you recommend your friend. Your friend is grateful and takes the job. On the first night, the child gets out of bed and, while your friend has fallen asleep watching television, drinks cleaning fluid from beneath the kitchen sink. The child is taken by an ambulance to the hospital and stays there for 2 days for observation and treatment. The married couple will not speak to you. Imagine yourself in such a situation and mark how likely you are to forgive your friend.

1   2   3   4   5

3. A friend offers to drop off a job application for you at the post office by the deadline for submission. A week later, you get a letter from the potential employer saying that your application could not be considered because it was postmarked after the deadline and they had a very strict policy about this. Your friend said that he or she met an old friend, went to lunch, and lost track of time. When he or she remembered the package, it was close to closing time at the post office and he or she would have to have rushed frantically to get there; he or she decided that deadlines usually aren't that strictly enforced so he or she waited until the next morning to deliver the package. Imagine yourself in such a situation and mark how likely you are to forgive your friend for not delivering the application on time.

1   2   3   4   5

4. You just started a new job and it turns out that a classmate from high school works there too. You think this is great; now you don't feel like such a stranger. Even though the classmate wasn't part of your crowd, there's at least a face you recognize. You two hit it off right away and talk about old times. A few weeks later, you are having lunch in the cafeteria and you overhear several of your coworkers, who do not realize you are nearby, talking about you and laughing; one even sounds snide and hostile toward you. You discover that your old classmate has told them about something you did back in school that you are deeply ashamed of and did not want anyone to know about. Imagine yourself in such a situation and mark how likely you are to forgive your old classmate for telling others your secret.

1   2   3   4   5

5. A distant cousin you haven't seen since childhood calls you one day and asks if he can stay with you while he looks for work and an apartment. You say it will be fine. He asks you to pick him up from the bus station that night and you do so. Your cousin is just like you fondly remember him; you reminisce for several hours. The next morning you give him some advice on job and apartment hunting in the area, then you go about your own business. That night you come home and witness an angry argument in front of your residence between your cousin and a

neighbor. Your cousin is obviously very drunk, cursing, and out of control. You ask what's happening and without really taking the time to recognize you, your cousin throws a bottle at you, cutting the side of your head. The police arrive and, with some scuffling, take your cousin away and take you to the emergency room where you have stitches put on your cut. The next afternoon, your cousin calls from the police station. He says he is really sorry about the whole scene and that it was not like him but he was upset about being turned down for three jobs that day. Imagine yourself in such a situation and mark how likely you are to forgive your cousin.

| 1 | 2 | 3 | 4 | 5 |

REFERENCES

Andersen, E. B. (1995). Polytomous Rasch models and their estimation. In G. H. Fischer & I. Molenar (Eds.), *Rasch models: Foundations, recent developments and applications* (pp. 271-291). New York: Springer-Verlag.

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika, 43,* 561-573.

Asendorpf, J. B. (1998). Personality effects on social relationships. *Journal of Personality and Social Psychology, 74,* 1531-1544.

Ashton, M. C., Paunonen, S. V., Helmes, E., & Jackson, D. N. (1998). Kin altruism, reciprocal altruism, and the Big Five personality factors. *Evolution and Human Behavior, 19,* 243-255.

Baumeister, R. F., & Cairns, K. J. (1992). Repression and self-presentation: When audiences interfere with self-deceptive strategies. *Journal of Personality and Social Psychology, 62,* 141-156.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117,* 497-529.

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

Buss, A. H., & Perry, M. (1992). The Aggression Questionnaire. *Journal of Personality and Social Psychology, 63,* 452-459.

Caprara, G. V. (1986). Indicators of aggression: The Dissipation-Rumination Scale. *Personality and Individual Differences, 7,* 763-769.

Caprara, G. V., Manzi, J., & Perugini, M. (1992). Investigating guilt in relation to emotionality and aggression. *Personality and Individual Differences, 13,* 519-532.

Collins, K., & Bell, R. (1997). Personality and aggression: The Dissipation-Rumination Scale. *Personality and Individual Differences, 22,* 751-755.

Crowne, D. P. (1979). *The experimental study of personality.* Hillsdale, NJ: Lawrence Erlbaum.

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24,* 349-354.

Davila, J., & Bradbury, T. N. (1998). Psychopathology and the marital dyad. In L. L'Abate (Ed.), *Handbook of family psychopathology.* New York: Guilford.

DiBlasio, F. A. (1998). The use of decision-based forgiveness intervention within intergenerational family therapy. *Journal of Family Therapy, 20,* 77-94.

Diehl, M., Elnick, A. B., Bourbeau, L. S., & Labouvie-Vief, G. (1998). Adult attachment styles: Their relations to family context and personality. *Journal of Personality and Social Psychology, 74,* 1656-1669.

Draba, R. E. (1977). *The identification and interpretation of item bias* (MESA Research Memoranda No. 25). Chicago: University of Chicago Press.

Dweck, C. S., Hong, Y., & Chiu, C. (1994). Implicit theories: Individual differences in the likelihood and meaning of dispositional inference. *Personality and Social Psychology Bulletin, 19,* 644-656.

Emmons, R. A. (2000). Personality and forgiveness. In M. E. McCullough, K. I. Pargament, & C. E. Thoresen (Eds.), *Forgiveness: Theory, research, and practice* (pp. 156-175). New York: Guilford.

Enright, R. D., Gassin, E. A., & Wu, C. (1992). Forgiveness: A developmental view. *Journal of Moral Education, 21,* 99-114.

Enright, R. D., & North, J. (Eds.). (1998). *Exploring forgiveness.* Madison: University of Wisconsin Press.

Fischer, G. H., & Molenar, I. (Eds.). (1995). *Rasch models: Foundations, recent developments and applications.* New York: Springer-Verlag.

Gonzales, M. H., Haugen, J. A., & Manning, D. J. (1994). Victims as "narrative critics": Factors influencing rejoinders and evaluative responses to offenders' accounts. *Personality and Social Psychology Bulletin, 20,* 691-704.

Gonzales, M. H., Manning, D. J., & Haugen, J. A. (1992). Explaining our sins: Factors influencing offender accounts and anticipated victim responses. *Journal of Personality and Social Psychology, 62,* 958-971.

Gorsuch, R. A., & Hao, J. Y. (1993). Forgiveness: An exploratory factor analysis and its relationship to religious variables. *Review of Religious Research, 34,* 333-347.

Graziano, W. G., Jensen-Campbell, L. A., & Hair, E. C. (1996). Perceiving interpersonal conflict and reacting to it: The case for agreeableness. *Journal of Personality and Social Psychology, 70,* 820-835.

Greenberg, M. A. (1995). Cognitive processing of traumas: The role of intrusive thoughts and reappraisals. *Journal of Applied Social Psychology, 25,* 1262-1296.

Hargrave, T. D., & Sells, J. N. (1997). The development of a forgiveness scale. *Journal of Marital and Family Therapy, 23,* 41-62.

Hodgins, H. S., Liebeskind, E., & Schwartz, W. (1996). Getting out of hot water: Facework in social predicaments. *Journal of Personality and Social Psychology, 71,* 300-314.

Huang, S. T., & Enright, R. D. (2000). Forgiveness and anger-related emotions in Taiwan: Implications for therapy. *Psychotherapy, 37,* 71-79.

John, O. P. (1990). The "Big Five" factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 66-100). New York: Guilford.

John, O. P., Donahue, E. M., & Kentle, R. (1992). *The Big-Five Inventory—Versions 4a and 54* (Tech. Rep.). Berkeley: University of California, Institute of Personality and Social Research.

Kaplan, B. H. (1992). Social health and the forgiving heart: The Type B story. *Journal of Behavioral Medicine, 15,* 3-14.

Larson, D. B., Swyers, J. P., & McCullough, M. E. (Eds.). (1998). *Scientific research on spirituality and health: A consensus report.* Baltimore: National Institute for Healthcare Research.

Levin, J. S., & Taylor, R. J. (1998). Panel analyses of religious involvement and well-being in African Americans: Contemporaneous vs. longitudinal effects. *Journal for the Scientific Study of Religion, 37,* 695-709.

Mauger, P. A., Perry, J. E., Freeman, T., Grove, D. C., McBride, A. G., & McKinney, K. E. (1992). The measurement of forgiveness: Preliminary research. *Journal of Psychology and Christianity, 11,* 170-180.

McAdams, D. P. (1994). *The person: An introduction to personality psychology* (2nd ed.). Fort Forth, TX: Harcourt, Brace.

McCullough, M. E. (2000). Forgiveness as human strength: Theory, measurement, and links to well-being. *Journal of Social and Clinical Psychology, 19,* 43-55.

McCullough, M. E., Exline, J. J., & Baumeister, R. F. (1998). An annotated bibliography of research on forgiveness and related concepts. In E. L. Worthington, Jr. (Ed.), *Dimensions of forgiveness: Psychological research and theological perspectives* (pp. 193-317). Philadelphia: Templeton Foundation Press.

McCullough, M. E., Pargament, K. I., & Thoresen, C. (Eds.). (2000). *The frontiers of forgiving: Conceptual, empirical, and clinical perspectives.* New York: Guilford.

McCullough, M. E., Rachal, K. C., & Hoyt, W. T. (2000). The assessment of forgiving. In M. E. McCullough, K. I. Pargament, & C. E. Thoresen (Eds.), *The frontiers of forgiveness: Theoretical, empirical, and clinical perspectives* (pp. 65-88). New York: Guilford.

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

McCullough, M. E., & Worthington, E. L., Jr. (2000). Religion and the forgiving personality. *Journal of Personality*, *67*, 1141-1164.

McCullough, M. E., Worthington, E. L., Jr., & Rachal, K. C. (1997). Interpersonal forgiving in close relationships. *Journal of Personality and Social Psychology*, *73*, 321-336.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen, Denmark: Danmarks Paedogogiske Institut.

Revelle, W. (1995). Personality processes. *Annual Review of Psychology*, *46*, 295-328.

Roberts, R. C. (1995). Forgivingness. *American Philosophical Quarterly*, *32*, 289-306.

Sandage, S. J., Worthington, E. L., Jr., Hight, T. L., & Berry, J. W. (2000). Seeking forgiveness: Theoretical context and an initial empirical study. *Journal of Psychology and Theology*, *28*, 21-35.

Schonbach, P. (1990). *Account episodes: The management and escalation of conflict*. Cambridge, UK: Cambridge University Press.

Smith, R. M. (1996). A comparison of methods for determining dimensionality in Rasch measurement. *Structural Equation Modeling*, *3*, 25-40.

Spielberger, C. D., Jacobs, G., Russell, S., & Crane, R. S. (1983). Assessment of anger: The State-Trait Anger Scale. In J. N. Butcher & C. D. Spielberger (Eds.), *Advances in personality assessment* (Vol. 2, pp. 161-189). Hillsdale, NJ: Lawrence Erlbaum.

Subkoviak, M. J., Enright, R. D., Wu, C. R., Gassin, E. A., Freedman, S., Olson, L. M., & Sarinopoulos, I. (1995). Measuring interpersonal forgiveness in late adolescence and middle adulthood. *Journal of Adolescence*, *18*, 641-655.

Tangney, J. P., Wagner, P., Barlow, D. H., & Marschall, D. E. (1996). Relation of shame and guilt to constructive vs. destructive responses to anger across the lifespan. *Journal of Personality and Social Psychology*, *70*, 797-809.

Thoresen, C. E., & Goldberg, J. H. (1998). Coronary heart disease: A psychosocial perspective on assessment and intervention. In S. Roth-Roemer, S.E.R. Kurpius, & C. Carmin (Eds.), *The emerging role of counseling psychology in health care* (pp. 94-136). New York: Norton.

Thoresen, C. E., Harris, A.H.S., & Luskin, F. (2000). Forgiveness and health: An unanswered question. In M. E. McCullough, K. I. Pargament, & C. E. Thoresen (Eds.), *Forgiveness: Theory, research, and practice* (pp. 254-280). New York: Guilford.

Thoresen, C. E., Luskin, F., & Harris, A.H.S. (1998). The science of forgiving interventions: Reflections and suggestions. In E. L. Worthington, Jr. (Ed.), *Dimensions of forgiveness: Psychological research and theological speculations* (pp. 163-190). Philadelphia: Templeton Foundation Press.

Van de Viert, E., & Euwema, M. C. (1994). Agreeableness and activeness as components of conflict behaviors. *Journal of Personality and Social Psychology*, *66*, 674-687.

Wade, S. H. (1989). *The development of a scale to measure forgiveness*. Unpublished doctoral dissertation, Fuller Theological Seminary, Pasadena, CA.

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, *92*, 548-573.

Weiner, B., Graham, S., Peter, O., & Zmuidinas, M. (1991). Public confession and forgiveness. *Journal of Personality*, *59*, 281-312.

Williams, R. (1989). *The trusting heart*. New York: Times Books.

Worthington, E. L., Jr. (1998a). *Dimensions of forgiveness: Psychological research and theological perspectives*. Philadelphia: Templeton Foundation Press.

Worthington, E. L. (Ed.). (1998b). The pyramid model of forgiveness: Some interdisciplinary speculations about unforgiveness and the promotion of forgiveness. In *Dimensions of forgiveness: Psychological research and theological perspectives* (pp. 29-55). Philadelphia: Templeton Foundation Press.

Worthington, E. L., Jr., Sandage, S. J., & Berry, J. W. (2000). Group interventions to promote forgiveness: What researchers and clinicians ought to know. In M. E. McCullough, K. I. Pargament, & C. E. Thoresen (Eds.), *Forgiveness: Theory, research, and practice* (pp. 228-254). New York: Guilford.

Worthington, E. L., Jr., & Wade, N. G. (1999). The social psychology of unforgiveness and forgiveness and implications for clinical practice. *Journal of Social and Clinical Psychology*, *18*, 385-418.

Wright, B. D. (1996). Sample size again. *Rasch Measurement Transactions*, *9*, 468.

Wright, B. D., & Douglas, G. A. (1975). *Better procedures for sample-free item analysis* (Research Memorandum No. 20). Chicago: University of Chicago, Statistical Laboratory, Department of Education.

Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square-fit values. *Rasch Measurement Transactions*, *8*, 370.