

The Voodoo Doll Task: Introducing and Validating a Novel Method for Studying Aggressive Inclinations

C. Nathan DeWall^{1*}, Eli J. Finkel², Nathaniel M. Lambert³, Erica B. Slotter⁴,
Galen V. Bodenhausen², Richard S. Pond Jr¹, Claire M. Renzetti⁵, and Frank D. Fincham⁶

¹Department of Psychology, University of Kentucky, Lexington, Kentucky

²Department of Psychology, Northwestern University, Evanston, Illinois

³Department of Family Sciences, Brigham Young University, Provo, Utah

⁴Department of Psychology, Villanova University, Villanova, Pennsylvania

⁵Department of Sociology and Center for Research on Violence Against Women, University of Kentucky, Lexington, Kentucky

⁶Family Institute, The Florida State University, Tallahassee, Florida

Aggression pervades modern life. To understand the root causes of aggression, researchers have developed several methods to assess aggressive inclinations. The current article introduces a new behavioral method—the *voodoo doll task* (VDT)—that offers a reliable and valid trait and state measure of aggressive inclinations across settings and relationship contexts. Drawing on theory and research on the law of similarity and magical beliefs (Rozin, Millman, & Nemeroff [1986], *Journal of Personality and Social Psychology*, 50, 703–712), we propose that people transfer characteristics of a person onto a voodoo doll representing that person. As a result, causing harm to a voodoo doll by stabbing it with pins may have important psychological similarities to causing actual harm to the person the voodoo doll represents. Nine methodologically diverse studies (total $N = 1,376$) showed that the VDT had strong reliability, construct validity, and convergent validity. Discussion centers on the importance of magical beliefs in understanding the causes of aggressive inclinations. *Aggr. Behav.* 9999:XX–XX, 2013. © 2013 Wiley Periodicals, Inc.

Keywords: aggression; violence; intimate partner violence; magical beliefs

INTRODUCTION

Modern social life thrives from cooperation and crumbles from conflict. Although aggression may have served an adaptive function several millennia ago (MacDonald, 2008), it is largely a destructive means of resolving conflict today. Yet aggression persists. Intimate partner violence occurs at alarmingly high rates (Schafer, Caetano, & Clark, 1998). In the 40 years after the end of World War II, there were roughly 150 wars and only 26 days of world peace (defined as the absence of international war) (Sluka, 1992). Civil wars rage on. And the threat of terrorist attacks pervades global consciousness.

Insight into the causes and consequences of aggression is facilitated by several factors, including the development of reliable and valid measures of aggressive inclinations. Aggression is defined as any behavior intended to cause harm to someone who is motivated to avoid the harm (Anderson & Bushman, 2002; Baron & Richardson, 1994; Bushman & Huesmann, 2010). Over its long history, aggression research has benefitted from

the development and validation of a broad range of measures to assess aggression. But how people perform aggression research continually evolves—and new measures must accommodate those changes. Specifically, aggression research now exists in at least three different research settings (i.e., one with no computers, one with a computer in a laboratory, and one with data collection taking place over the Internet) and in at least three different relationship contexts (i.e., between strangers,

Contract grant sponsor: National Science Foundation; contract grant numbers: BCS-719780, BCS-1104118; contract grant sponsor: University of Kentucky College of Arts and Sciences.

C. Nathan DeWall and Eli J. Finkel contributed equally to this work.

*Correspondence to: C. Nathan DeWall, Department of Psychology, University of Kentucky, Lexington, KY 40506-0044.

E-mail: nathan.dewall@uky.edu

Received 21 November 2011; Revised 30 April 2013; Accepted 3 June 2013

DOI: 10.1002/ab.21496

Published online XX Month Year in Wiley Online Library (wileyonlinelibrary.com).

between nonromantic members of one's social network, and between romantic partners). What worked 40 years ago to assess aggression may still have validity, but it may not be equally valid across all settings or all relationship contexts.

The current investigation introduces a new behavioral method for measuring aggressive inclinations that has strong validity in all of these research settings and across all of these relationship contexts—the *voodoo doll task* (VDT). The VDT is not meant to supersede existing aggression measures; rather, it aims to complement existing measures by accommodating the multiple settings and relationship contexts aggression researchers now use. Drawing on research on magical beliefs and the law of similarity (Rozin, Millman, & Nemeroff, 1986), we propose that people transfer characteristics of a person onto a voodoo doll representing that person. As a result, the process of causing harm to a voodoo doll by stabbing it with pins has important psychological similarities to the process of causing actual harm to the person the voodoo doll represents. The current investigation offers converging evidence that the VDT represents a flexible method for assessing aggressive inclinations in multiple settings and relationship contexts. Nine methodologically diverse studies showcase the VDT's reliability and validity as both a trait and a state measure of aggressive inclinations.

The first section of this article discusses common methods for measuring aggressive inclinations toward strangers and close relationship partners. It also discusses the potential limitations of these methods in terms of their use across research settings and relationship contexts. The second section presents the theoretical rationale underlying the VDT and explains how it has strong validity across settings and relationship contexts. The third section presents findings from nine methodologically diverse studies that demonstrate the VDT's reliability and validity. The fourth section discusses the limitations of the VDT and directions for future research. The fifth, concluding section discusses the usefulness of the VDT as a measure of aggressive inclinations in multiple settings and relationship contexts.

WIDELY USED MEASURES OF AGGRESSIVE INCLINATIONS TOWARD STRANGERS AND CLOSE RELATIONSHIP PARTNERS

Scholars have developed many methods to measure aggressive inclinations. These methods have enabled researchers to gain insight into the causes and, on occasion, the consequences of aggressive inclinations inside and outside the laboratory. But no method is perfect. In what follows, we provide a brief summary of commonly used methods to assess aggressive tendencies toward strangers and close relationship partners. Each

method has passed tests of reliability and validity, which is demonstrated in the validation papers we cite below. We also note some limitations of each method, focusing on its use in different settings and relationship contexts. Our list is not meant to be an exhaustive summary of all methods used to assess aggressive inclinations; it simply covers the most widely used measures.

Taylor Aggression Paradigm and Its Variants

A widely used and well-validated method for measuring aggression toward strangers is the Taylor Aggression Paradigm (TAP; Taylor, 1967). In the TAP, participants complete a competitive reaction-time game against an opponent who is either real or fictitious (i.e., a computer program mimics an opponent's responses). Whoever responds first wins the trial and gets the opportunity to administer a noxious stimulus to the loser. In the initial version of the TAP (Taylor, 1967), the noxious stimulus consists of electric shocks. In the most common variant of the TAP, noxious stimulus consists of noise blasts (Bushman & Baumeister, 1998). Typically, participants select the intensity and duration levels of the noxious stimuli, which can be used separately or in combination to create various composites of aggressive behavior. A nonaggressive option is also frequently included, in which participants can select no noxious stimulus for their opponent.

The TAP has many strengths, but it also has setting and relationship context limitations. It can be used in only one setting, namely in a laboratory with a computer. It would be difficult to use the TAP without a computer (or some other complex apparatus), and it would be nearly impossible to use it in many of the most common forms of Internet-based research scholars typically employ. In addition, the TAP may not be readily applicable to close relationship contexts. If people know that their friend or close relationship partner is extremely unaggressive or has comparably faster reaction-times, they may find it implausible that their partner administered a high intensity noxious stimulus or was slower to respond on the task. Furthermore, if people have a playful but competitive interaction style with their friend or partner, TAP behavior might not accurately represent aggression in a theoretically precise way.

Hot Sauce Paradigm

The hot sauce paradigm has gained widespread use in recent years (Lieberman, Solomon, Greenberg, & McGregor, 1999). In this innovative paradigm, participants are led to believe that they will prepare a food sample for another person. By way of a rigged drawing, participants are told that they will prepare a food sample that contains hot and spicy ingredients. Participants are then given a sheet of paper that lists the person's food

preferences and allergies, which make it apparent that the person dislikes hot and spicy foods. The participant is then given a container of hot sauce and is instructed to give as much hot sauce as he or she wishes on a piece of food, with the proviso that the other person will have to consume the entire food sample. The amount of hot sauce the participant gives the other participant is used as the measure of aggression.

There are some limitations to the hot sauce paradigm. For example, it is ideally suited for measuring aggression in one setting, namely without a computer. It is not straightforward to dole out hot sauce using a laboratory computer or over the Internet. The hot sauce paradigm is also not ideally suited for measuring aggression toward non-strangers. If people know that a close relationship partner loves hot and spicy foods, increased allocation of hot sauce to that person might be better conceptualized in terms of considerateness than aggressiveness.

The Physical Assault and Psychological Aggression Subscales of the Conflict Tactics Scales

The gold standard measures of physical and psychological aggression between close relationship partners are the Physical Assault and Psychological Aggression subscales of the Conflict Tactics Scales (CTS; Straus, 1979), the original version of which has been used in over 400 papers involving more than 70,000 participants (Archer, 2000; Straus, 1995). Updated in 1996 (Straus, Hamby, Boney-McCoy, & Sugarman, 1996), the Physical Assault subscale of the Revised Conflict Tactics Scales (CTS2) includes five items measuring “minor” aggression (e.g., “pushed or shoved my partner,” “slapped my partner”) and seven items measuring “severe” aggression (e.g., “beat up my partner,” “choked my partner”). The Psychological Aggression subscale includes four items measuring “minor” aggression (e.g., “insulted or swore at my partner,” “shouted or yelled at my partner”) and four items measuring “severe” aggression (e.g., “called my partner fat or ugly,” “accused my partner of being a lousy lover”).

In a typical study, participants report how often they did each of these behaviors in a previous period of time (e.g., over the past year) using this scale: 1 = *Once*, 2 = *Twice*, 3 = *3–5 times*, 4 = *6–10 times*, 5 = *11–20 times*, 6 = *more than 20 times*, 7 = *not once in the past year* [or other specified time period], *but it did happen before*, and 0 = *this has never happened*. In some studies, participants also report how often their close relationship partner did these things to them. Responses are used to create a composite measure of physical assault tendencies. Discussion of the strengths of the CTS and CTS2 can be found in Straus et al. (1996; see also Archer, 2000).

The CTS and CTS2’s main limitation rests with the relationship context being assessed. An ongoing relationship is required for participants to complete the CTS and CTS2. It is not straightforward to use these scales to assess aggression between people who have just met.

Intimate Partner Violence Inclinations

In addition to self-report measures of actual aggressive behavior, researchers use self-report measures that assess aggressive inclinations. Two such measures are the Proximal Antecedents to Violence Episodes Scale (PAVE; Babcock, Costa, Green, & Eckhardt, 2004) and the Articulated Thoughts in Simulated Situations (ATSS; Eckhardt, Barbour, & Davison, 1998). Both measures present participants with situations in which one’s partner provokes them and situations in which one’s partner does not. Participants then either report the likelihood that they would behave with physical aggression toward their partner (using the PAVE) or verbalize their response to the situation. Responses to the ATSS are then coded for remarks related to physical aggression (“If he ever did that to me, I would beat his ass;” Slotter et al., 2012). Responses to both measures are combined to create composite measures of aggressive inclinations. These measures have the same limitation of the CTS and CTS2: They generally require an ongoing relationship between the participant and another person.

Summary and Goals for a New Measure of Aggressive Inclinations

This section reviewed several methods to assess aggressive responses to strangers and close relationship partners. Each method has strengths and weaknesses in terms of research setting and relationship type. Based on this review, a new measure of aggressive inclinations should address the following two limitations. First, it should be amenable for use in the three main research “worlds:” settings without a computer, laboratories with computers, and over the Internet. Second, it should be flexible enough to assess aggressive inclinations toward strangers, friends, and close relationship partners. The next section describes the VDT, which is a method for assessing aggressive inclinations that addresses these limitations. To be clear, we are not suggesting that the VDT is superior to other aggressive inclinations in terms of reliability and validity; it is simply a measure of aggressive inclinations that addresses limitations of research setting and relationship type inherent in existing measures.

THE VOODOO DOLL TASK

The VDT offers participants the opportunity to inflict harm on a doll that represents another person by stabbing

the doll with pins. In a typical study, participants complete individual difference measures related to aggression or are exposed to a situational manipulation designed to stimulate an aggressive urge. Participants are then shown a doll that represents a close relationship partner or a person involved in the situational manipulation (e.g., a person who insulted, rejected, or otherwise provoked the participant). The voodoo doll is presented either by handing participants an actual doll or by presenting them with a computer-based image of a doll. Typically, participants are then told that because some parts of the study may have been negative, they will be given the opportunity to release their negative energy by stabbing the doll with as many pins as they wish. The instructions for the VDT do not use the word “voodoo” at any time, and instructions regarding the release of negative energy are not required. Stabbing the doll with more pins indicates higher levels of aggressive intentions.

Therefore, the VDT addresses the two common limitations with aggression measures. First, it can be used in any research setting. If researchers do not wish to use a computer, they can hand a participant a doll with some straight pins. If researchers desire to use a computer to administer the task, they can do so in the laboratory or over the Internet. Second, the VDT can be used to assess aggressive inclinations toward a person from any relationship context. All that is needed is to adjust the instructions to ensure participants that the doll represents the intended target.

These strengths notwithstanding, the VDT also has several potential limitations. We discuss the most important potential limitation here: Is it plausible that people, especially educated people, view the voodoo doll as a proxy for the target of aggression? If so, then people are likely to insert more pins into the voodoo doll precisely under those circumstances when they are likely to enact other forms of aggression, and the VDT can potentially serve as a useful measure of aggression. If not, then people are *unlikely* to insert more pins into the voodoo doll precisely under those circumstances when they are likely to enact other forms of aggression, and the VDT cannot serve as a useful measure of aggressive inclinations.

Do people Take Voodoo Dolls Seriously?

The construct validity of the VDT hinges on whether people take seriously the harm they cause to images representing other people. Do people take voodoo dolls seriously? A growing body of theoretical and empirical work suggests that they do.

In an influential paper, Rozin et al. (1986) proposed that although magical beliefs are stereotypically associated with economically underdeveloped, traditional

cultures, members of contemporary American society are also routinely susceptible to such beliefs. In one study, participants threw darts at the images of faces of five people: John F. Kennedy, Adolph Hitler, the experimenter, a strongly liked person, and a strongly disliked person (these last two faces were selected idiographically by each participant). As predicted, participants were significantly less accurate than usual when aiming at Kennedy or the strongly liked person, but not when aiming at any of the other targets, suggesting that they were uncomfortable piercing the images of the face of people they like or admire. The expected reversal for Hitler and the strongly disliked person did not emerge, most likely due to ceiling effects: “Subjects’ motivation to perform on the dart task was high, and they were probably throwing as well as they could at a blank target, so that an enhancement with a negative target may not have been possible” (Rozin et al., 1986, p. 709). This study was a landmark in the study of *magical thinking*, which refers to beliefs that an agent can influence events at a distance in the absence of any physical explanation (Eckblad & Chapman, 1983; Nemeroff & Rozin, 2000; Pronin, Wegner, McCarthy, & Rodriguez, 2006; Woolley, 1997; Zusne & Jones, 1989).

Rozin et al. (1986) argued that these results emerged because of the operation of the magical law of similarity, which holds that *image = object*; that is, the image of something (whether a mental image or a physical representation, such as a photo or a voodoo doll) is equated with the thing itself, and anything that happens to the image will affect the depicted object too. For example, people are eager to avoid throwing a dart at the face of a beloved person, presumably due to a gut-level intuition that the dart might actually harm the person. Similarly, people are especially resistant to throw a dart at an image of someone who is incapable of behaving aggressively, such as a baby (King, Burton, Hicks, & Drigotas, 2007).

Of course, such a concern is not likely to be consciously, explicitly endorsed by most educated adults, who are unlikely to acknowledge any belief in magic. However, our minds are populated not only by rational beliefs but also by more primitive responses. Superstitions operate primarily via automatic and relatively nonconscious means, rather than by explicit belief (Kramer & Block, 2008). The philosopher Gendler (2008) proposed the term “alief” (in contrast to “belief”) as a label for primitive reactions that arise intuitively and that can often be at odds with our conscious, deliberated beliefs. Notably, she argued that aliefs are affect-laden and action-generating. Aliefs and beliefs often compete for control over behavior, and, not uncommonly, the alief wins (see also Strack & Deutsch, 2004). This tension is evident in many of Rozin et al. (1986) experiments, in

which magical thinking leads to behavior that is obviously illogical yet also hard to resist.

Why would people have these magical intuitions? McKay and Dennett (2009) review a range of evolutionary mechanisms that can account for the emergence of “misbeliefs.” One integrative theory of magical and superstitious thinking was proposed by Lindeman and Aarnio (2007). They begin by arguing, in line with a vast literature on infant cognition, that children possess core knowledge (i.e., knowledge that is not acquired from experience) in the physical, psychological, and biological domains. In their view, magical thinking reflects the inadvertent blending of innate, intuitive knowledge across these distinct domains, as when non-intentional events from the physical domain are conflated with intentional events from the psychological domain. For example, people may interpret an avalanche as having a motivation to attack mountaineers because it caused harm to people who were motivated to avoid the harm. Whatever their origins, holding and acting on superstitious beliefs can actually be productive (Damisch, Stoberock, & Mussweiler, 2010; Ng, Chong, & Du, 2010), thereby reinforcing their seeming validity and likelihood of future deployment.

When it comes to the specific matter of stabbing a voodoo doll, although most educated people are unlikely to explicitly endorse the idea that this action will harm the person, they may nevertheless be subject to a primitive feeling that something bad might result. Moreover, stabbing the doll could be seen as an instance of “tempting fate,” and people do believe that tempting fate increases the probability of negative events (Risen & Gilovich, 2008). Consequently, they are reluctant to harm the image of another person—especially when they want to avoid harming the person, as in the “beloved person” condition examined by Rozin et al. (1986). Along similar lines, Pronin et al. (2006) showed that people could be led to believe that they had actually harmed another person via a voodoo hex, particularly when they had previously been experimentally induced to harbor ill will toward the person.

Wanting to harm someone thus appears to enhance the plausibility of magically harming them. For these reasons, stabbing a voodoo doll provides an appealing way to assess aggressive inclinations that carries a primitive kind of psychological reality, albeit in an ethically acceptable way. Crucially, Denzler, Förster, and Liberman (2009) showed that stabbing a voodoo doll satisfied the goal of retaliating against a provocation. Specifically, they asked participants to imagine their best friend making advances toward the participant’s own romantic partner. Some participants subsequently had the opportunity to stab a voodoo doll representing the best friend/perpetrator, whereas others did not. Quite sensibly,

the provocation resulted in heightened accessibility of aggression-related thoughts. This heightened accessibility of aggression-related thoughts relates to the Zeigarnik effect (Zeigarnik, 1927), in which people experience intrusive thoughts when they start activities but do not finish them. Having the opportunity to stab the voodoo doll resulted in subsequently decreased accessibility of these thoughts, thereby reducing the Zeigarnik effect. But aggressive thoughts remained highly accessible among the participants who had no opportunity to act on their aggressive impulses. In other words, stabbing a voodoo doll appeared to satisfy the goal of aggressing against an offender.

Thus, there is theoretical and empirical precedent for the use of the voodoo doll as a measure of aggressive inclinations. Given the law of similarity (Rozin et al., 1986), participants should project another person’s characteristics onto a doll meant to represent that other person.

Current Research

In nine studies, the current investigation introduced and validated the VDT as a new measure of aggressive inclinations with equal validity across different research settings and relationship contexts. Participants in all studies were presented with a voodoo doll and were instructed to insert into it as many pins as they wished. Consistent with the law of similarity, we predicted that participants would transfer characteristics of the other person onto the voodoo doll. Consequently, personality traits and situational factors that predispose people to behave aggressively should be associated with inserting more pins into the voodoo doll.

We sought to establish the reliability of voodoo doll behavior in two ways. First, we correlated voodoo doll stabbing behavior between two time points. Second, we measured the internal reliability of participants’ daily pin insertion.

We also sought to verify the construct validity and convergent validity of the VDT. First, we showed that the VDT has strong construct validity by demonstrating that VDT responses relate to established measures of aggressive tendencies. We showed that the VDT has strong convergent validity by demonstrating that responses are associated with a variety of self-report indicators of aggression, such as trait physical aggression and both physical assault and psychological aggression perpetrated against a close relationship partner. In addition, we demonstrated that provocation, a situational factor that commonly increases aggression, also increases the number of pins that participants insert into the voodoo doll. We also showed that pin insertion is associated with other behavioral measures of aggressive inclinations, such as calling one’s partner names during a

problem solving-task, behaving angrily during a conflict discussion task, and blasting a close relationship partner with intense and prolonged bursts of white noise in the TAP.

STUDY 1: ASSOCIATIONS WITH DISPOSITIONAL PHYSICAL AGGRESSIVENESS AND IPV INCLINATIONS

Study 1 sought to provide initial evidence that VDT responses correlate with dispositional physical aggressiveness and inclinations toward perpetrating intimate partner violence (IPV). We predicted that participants with higher scores on measures of both dispositional tendencies toward physical aggressiveness and inclinations toward IPV would insert more pins into the doll.

Method

Participants. Eighty one undergraduates (63 women, 18 men¹) participated in this study in exchange for partial course credit. Of these participants, the 39 who were currently involved in a romantic relationship reported on that relationship, and the 42 who were not reported on a close friend. Participants were not randomly assigned to report on either a romantic partner or a close friend. Results did not differ across these two groups, which confirms our prediction that the VDT would have comparable validity across different relationship contexts. Therefore, responses were collapsed across partner context. Age was not assessed in this study, but participants came from an undergraduate class in which the average age is approximately 19 years old.

Measures

VDT. Participants were shown a picture of a doll on a computer that they were told represented their current (or, for those who were not romantically involved, their most recent) romantic partner or a close friend. Participants were told that they could release any negative energy they experienced during the study by inserting pins into the voodoo doll. Participants could select from 0 to 51 pins. The average number of pins inserted into the doll was 2.54 ($SD = 7.23$). Overall, 63.0% of participants did not insert any pins, 33.8% inserted 1–10 pins, and 3.2%

inserted more than 10 pins. Because each response on the VDT represents a discrete event (i.e., a pin), we specified a Poisson distribution; such a distribution provides estimates of the number of events that occur in a fixed time period (i.e., number of pins inserted over the course of a study) (Gardner, Mulvey, & Shaw, 1995). To supplement these analyses, we conducted additional analyses using negative binomial regression analysis. This type of analysis is useful when the ratio of the standard deviation to the mean is greater than one, which represents what is called “overdispersion” (Long, 1997).

Physical Aggressiveness and IPV measures.

Participants completed Buss and Perry’s (1992) widely used nine-item measure of dispositional tendencies toward *physical aggressiveness*. Sample items were, “Once in a while, I can’t control the urge to strike another person,” “I can think of no good reason for ever hitting a person” (reverse-scored), and “If I have to resort to violence to protect my rights, I will” (1 = *Extremely Uncharacteristic of me*, 7 = *Extremely Characteristic of me*; $\alpha = .87$). Not surprisingly, higher scores on this measure relate to stronger tendencies toward physical aggression (Giancola & Parrott, 2008). Participants also completed the Babcock et al. (2004) 20-item measure of *IPV perpetration inclinations*. This scale, which we modified for dating rather than married samples (Finkel, DeWall, Slotter, Oaten, & Foshee, 2009), had participants indicate how likely they would become physically aggressive in response to each of 20 hypothetical provocative partner behaviors (1 = *Not at all likely that I would be physically aggressive*, 9 = *Extremely likely that I would be physically aggressive*; $\alpha = .96$). Sample items were, “My partner ridicules or makes fun of me,” “My partner does something to offend or ‘disrespect’ me,” and “I find out that my partner has been flirting with someone.” Responses on this measure did not differ as a function of relationship status (current romantic partner or close friend) and therefore results were collapsed across all participants.

Procedure. Participants completed the physical aggressiveness measures and then the VDT. Participants completed all components of the study over the Internet.

Results and Discussion

Our main hypothesis was that higher scores on the standardized measures of physical aggressiveness would relate to the insertion of a larger number of pins in the VDT. To test this hypothesis, we conducted two Poisson regression analyses predicting pin insertion, one including dispositional physical aggressiveness as the predictor and the other including IPV perpetration tendencies as the predictor. As expected, both general physical aggressiveness and IPV perpetration tendencies were significantly associated with pin insertion. These

¹In this and all other studies, we tested whether we observed gender interactions for VDT responses. No reliable interactions emerged across all nine studies. In terms of interactions, three significant interactions emerged of the 60 analyses we conducted (5%), all in the direction of men inserting more pins into the doll as a function of another predictor variable (Study 2: Minor physical assault and minor psychological aggression; Study 7: Major physical assault-1 year and Major physical assault-5 year). All other interactions were non-significant. Of the 60 analyses we conducted, 52 (87%) remained significant and 54 (90%) remained significant or marginally significant after controlling for gender. Thus, the VDT offers a method for assessing aggressive inclinations among both men and women.

TABLE I. Voodoo Doll Task Responses Correlate With Trait Physical Aggression and Intimate Partner Violence Perpetration Tendencies (Study 1)

Variable	<i>B</i>	χ^2	<i>df</i>	<i>p</i> -Value	<i>B</i>	χ^2	<i>df</i>	<i>p</i> -Value
Study 1	Poisson				Negative Binomial			
Trait Physical Aggression	0.34	4.75	1, 79	<.03	0.46	8.81	1, 79	.003
IPV perpetration tendencies	0.36	7.71	1, 78	.005	0.48	7.98	1, 78	.005
Controlling for Gender								
Trait Physical Aggression	0.28	1.88	1, 78	.17	0.42	3.77	1, 78	.052
IPV perpetration tendencies	0.35	7.72	1, 77	<.005	0.47	10.01	1, 77	.002

associations remained significant after controlling for relationship type (both $ps < .05$). Next, we conducted additional analyses using negative binomial regression analysis, which yielded similar results (see Table I).

In summary, these results employed self-report measures of general physical aggressiveness and IPV perpetration inclinations as a demonstration of the construct validity of the VDT. Participants with stronger tendencies toward aggression (in general or toward a romantic partner) inserted more pins into a voodoo doll representing their partner.

STUDY 2: AN ASSESSMENT OF TEST-RETEST RELIABILITY AND ASSOCIATIONS WITH THE CTS2

Study 2 had three goals. First, it sought to establish the test–retest reliability of responses to the VDT. Participants completed the VDT twice, separated by four weeks. Second, it sought additional evidence for the construct validity of the VDT. Participants completed the revised Conflict Tactics Scales (CTS2; Straus et al., 1996) measures of minor physical assault and psychological aggression toward their current romantic partner over the previous 8 weeks. Third, it sought to show that responses to the VDT were unrelated to social desirability.

Method

Participants. One hundred and forty undergraduates (111 women, 29 men) participated in this study in exchange for partial course credit. Average age was 20.99 ($SD = 2.42$). For these individuals, relationship length was: 45.1% 2+ years, 11.3% 19–24 months, 8.5% 13–18 months, 22.5% 7–12 months, 8.5% 4–6 months, 0.7% 2–3 months, and 3.5% less than 2 months. Participants completed the self-report and voodoo task measures in the laboratory initially and then once again four weeks later, reporting on their close friend at both occasions. Eighteen participants discontinued the study after time 1.

Measures

VDT. Participants completed the VDT using the same instructions as in Study 1. The average number of

pins inserted into the doll at time 1 was 1.41 ($SD = 3.93$) and at time 2 was 1.30 ($SD = 3.13$). At time 1, 68.5% of participants did not insert any pins, 28.7% inserted 1–10 pins, and 2.8% inserted more than 10 pins. At time 2, 71.8% of participants did not insert any pins, 25.8% inserted 1–10 pins, and 2.4% inserted more than 10 pins.

Convergent Validity Measures. Participants completed Straus and colleagues' (1996) measures of *minor physical assault* and *minor psychological aggression*, both of which were discussed in the Introduction. These measures assessed how often participants had behaved aggressively toward their partner over the previous eight weeks. Responses were summed to create composite physical assault and psychological aggression indices. Participants also completed a brief, previously validated version of the Marlowe–Crowne Social Desirability Scale (Strahan & Gerbasi, 1972) to assess the tendency to present oneself in a socially desirable manner.

Procedure. Participants completed all components of the study online. They completed the physical assault scale, the psychological aggression scale, and the VDT twice, four weeks apart.

Results and Discussion

First, we sought to establish the test-retest reliability of pin insertion in the VDT. As expected, pin insertion at time 1 correlated strongly with responses at time 2, $r = .58$, $p < .001$.² This correlation is strong (Cohen, 1988), especially for a single-item assessment instrument. The strong correlation between the two time points suggests that the VDT can capture stable aggressive tendencies (in the absence of any strong situational triggers of aggression).

Next, to test our primary hypothesis, we conducted four Poisson regression analyses using concurrent pin

²To ensure that this correlation was not affected by the non-normal nature of our dependent variable, we repeated this analysis using both Poisson and negative binomial regression. Both regressions yielded nearly identical results [Poisson: $\chi^2(1, 122) = 23.67$, $p < .001$; Negative Binomial: $\chi^2(1, 122) = 9.58$, $p < .01$].

insertion at the two time points as the dependent measures, two including physical assault as the predictor and two including psychological aggression as the predictor. As expected, both physical assault and psychological aggression exhibited significant cross-sectional associations with pin insertion at both time points, which remained significant after controlling for gender and social desirability. Similar effects emerged using negative binomial regression. Social desirability did not moderate any of the above associations (or relate to VDT responses at either time point; all p values $>.05$). The results are presented in Table II.

In summary, these results demonstrated strong test-retest reliability of pin insertion in the VDT. In addition, they employed self-report measures of physical assault and psychological aggression toward a current romantic partner, providing further evidence for the construct validity of the VDT. Participants with stronger tendencies toward physical assault and psychological aggression inserted more pins into a voodoo doll representing their partner. Finally, they demonstrated that the VDT assesses aggressive tendencies in a manner that appears unrelated to social desirability concerns.

STUDY 3: ASSOCIATIONS WITH LABORATORY-BASED PSYCHOLOGICAL AGGRESSION

Studies 1 and 2 demonstrated that established measures of general aggression (Buss & Perry, 1992), intimate partner violence (Babcock et al., 2004; Straus et al., 1996), and psychological aggression (Straus

et al., 1996) were associated with pin insertion in the VDT. Studies 3–5 adopted a different approach: rather than examining associations of existing self-report measures of aggression with pin insertion, these studies examined associations of laboratory-based behavioral measures of aggression with pin insertion. Study 3 examined psychological aggression (name-calling behavior during a dyadic task) as the laboratory-based behavioral measure.

Method

Participants. Forty-seven undergraduates (24 female) participated in the study for partial credit. Average age was 20.90 ($SD = 2.39$). Participants reported on their relationship with a romantic partner. Relationship length was: 27.1% 2+ years, 6.2% 19–24 months, 8.3% 13–18 months, 20.8% 7–12 months, 18.8% 4–6 months, 6.2% 2–3 months, and 12.5% less than 2 months.

Measures

VDT. Participants completed the VDT using the same instructions as in Studies 1 and 2. Participants were instructed that the voodoo doll represented their current romantic partner. The average number of pins inserted into the doll was 1.14 ($SD = 2.30$). In this sample, 61.9% of participants did not insert any pins, 38.1% inserted 1–10 pins, and no participants inserted more than 10 pins.

Procedure. Participants first completed the voodoo doll measure in a session. Four weeks later, they returned to the laboratory with their romantic partner. The experimenter informed them that the two of them would be videotaped as they completed a task together. The

TABLE II. Voodoo Doll Task Responses at Two Time Points Correlate With Physical Assault and Psychological Aggression (Study 2)

Variable	B	χ^2	df	p -Value	B	χ^2	df	p -Value
Study 2		Poisson			Negative Binomial			
VDT: Time 1								
Physical Assault	0.41	5.54	1, 138	<.05	0.40	5.07	1, 138	<.05
Psychological Aggression	0.37	4.96	1, 138	<.05	0.37	4.68	1, 138	<.05
VDT: Time 2								
Physical Assault	0.32	5.54	1, 125	<.05	0.34	6.29	1, 125	.01
Psychological Aggression	0.33	4.18	1, 125	<.05	0.36	5.16	1, 125	<.05
Controlling for Gender								
VDT: Time 1								
Physical Assault	0.44	5.52	1, 136	<.05	0.40	5.25	1, 136	<.05
Psychological Aggression	0.36	5.01	1, 136	<.05	0.38	4.83	1, 136	<.05
VDT: Time 2								
Physical Assault	0.32	6.03	1, 120	<.05	0.36	7.01	1, 120	<.01
Psychological Aggression	0.34	4.43	1, 120	<.05	0.37	5.52	1, 120	<.05
Controlling for Social Desirability								
VDT: Time 1								
Physical Assault	0.42	6.40	1, 130	<.05	0.44	5.94	1, 130	<.05
Psychological Aggression	0.41	6.36	1, 130	<.05	0.42	5.89	1, 130	<.05
VDT: Time 2								
Physical Assault	0.31	4.66	1, 119	<.05	0.39	5.96	1, 119	<.05
Psychological Aggression	0.48	6.73	1, 114	<.05	0.37	4.93	1, 114	<.05

experimenter then blindfolded the partner and instructed him or her to add details to a picture of a house, a tree, and two people (e.g., door on the house, apples on the tree). The experimenter instructed participants, who were not blindfolded, to provide instructions to their blindfolded partner.

Five trained coders subsequently watched each task on videotape. They used the following item to rate how often participants engaged in psychologically aggressive behavior during the problem-solving task: “How often did they call their partner names?” [1 = *never*; 5 = *constantly*; ICC = .71]. Some examples of names that were used were “jerk” and “jackass.”

We developed this task because it placed participants and their partner into a highly interdependent situation that could pull for motivations and emotions ranging from benevolence to frustration. We focused on name-calling because of its central role in operationalizations of psychological aggression. For example, the psychological aggression subscale of the CTS2 (Straus et al., 1996) includes items like, “insulted or swore at my partner,” “called my partner fat or ugly,” and “accused my partner of being a lousy lover.” Although the level of conflict in our laboratory-based problem-solving task—and the relevance of physical appearance or sexual ability—renders some of the CTS2 items less relevant, it is clear that name-calling is a major form of psychological aggression.

Results and Discussion

To test our hypothesis, we conducted a Poisson regression analysis that correlated participants’ pin insertion from the degree to which they called their partner names during the laboratory problem-solving task. As expected, higher levels of name-calling were significantly associated with pin insertion in the VDT, which remained significant when controlling for gender. Additional analyses using negative binomial regression yielded identical results (Table III).

This finding complements the self-report procedures from Studies 1 and 2 with behavioral evidence of the construct validity of the VDT. Participants with stronger tendencies toward psychological aggression on the dyadic problem-solving task inserted more pins into a voodoo doll representing their partner.

STUDY 4: ASSOCIATIONS WITH LABORATORY-BASED PHYSICAL AGGRESSION

Whereas Study 3 employed laboratory procedures to link pin insertion to a behavioral indicator of *psychological* aggression toward one’s romantic partner, Study 4 employed laboratory procedures to link pin insertion to a behavioral indicator of *physical* aggression toward one’s

close friend. Participants once again completed the VDT (representing the voodoo doll as a close friend) and, four weeks later, arrived at the laboratory with their partner. Using the same time interval as Study 2 enable us to examine the similarity in test–retest reliability coefficients across independent samples. After completing the VDT for the second time, they completed a noise-blast version of the TAP against their close friend. We predicted that participants who blasted their friend with more extreme noise on the TAP would insert a larger number of pins in the VDT.

Method

Participants. Two hundred ten undergraduates (144 women, 43 men, 23 did not report their gender) participated in this study in exchange for partial course credit. Average age was 19.44 ($SD = 1.76$). Participants completed the competitive reaction-time task against a close friend.

Measures

VDT. Participants completed the VDT using the same instructions as in the previous studies, with the target of the VDT being their close friend. The average number of pins inserted into the doll at each time point were 2.50 ($SD = 7.09$) and 2.71 ($SD = 8.52$), respectively. At time 1, 48.6% of participants did not insert any pins, 47.1% inserted 1–10 pins, and 4.3% inserted more than 10 pins. At time 2, 55.5% of participants did not insert any pins, 38.6% inserted 1–10 pins, and 5.9% inserted more than 10 pins.

TAP. Participants engaged in a version of the TAP in which they ostensibly competed against their close friend to see who could respond more quickly, with the winner blasting the other person with noise. Participants controlled the intensity of the noise (0–105 dB, about the same volume as a passing train) and how long their partner endured the noise (0–2.5 sec). This task is a well validated measure of laboratory aggression (e.g., Anderson & Bushman, 1997; Bushman & Baumeister, 1998; Giancola & Chermack, 1998). In reality, participants completed the reaction-time task against the computer, which was programmed to mimic another person’s actions. Of the 25 trials, the participant lost 13 (randomly determined). Intensity and duration levels of the opponent’s noise blasts toward the participant increased subtly across the course of the task. We programmed the aggression task in this way to mimic the violence escalation cycle in “real-life” situations (Anderson, Buckley, & Carnagey, 2008). With the noise blasts, participants controlled a weapon that, within the ethical limits of the laboratory, could be used to blast their partner with aversive and prolonged noise.

To provide convergent validity, we used three different measures of aggression from the competitive reaction-

TABLE III. Voodoo Doll Task Responses Correlate With Psychological Aggression (Study 3) and Physical Aggression (Study 4)

Variable	<i>B</i>	χ^2	<i>df</i>	<i>p</i> -Value	<i>B</i>	χ^2	<i>df</i>	<i>p</i> -Value
Study 3					Negative Binomial			
Controlling for Gender								
Name-Calling	2.40	10.82	1, 46	<.001	2.47	8.56	1, 46	<.01
Name-Calling	2.65	9.95	1, 44	<.01	2.58	8.51	1, 44	<.01
Study 4								
VDT: Time 1								
Unprovoked Aggression	0.34	14.25	1, 208	<.001	0.36	25.34	1, 208	<.001
Total Aggression	0.32	13.88	1, 208	<.001	0.34	20.47	1, 208	<.001
Extreme Aggression	0.07	13.89	1, 208	<.001	0.08	15.56	1, 208	<.001
VDT: Time 2								
Unprovoked Aggression	0.31	8.75	1, 208	.003	0.33	9.09	1, 208	.003
Total Aggression	0.23	4.53	1, 208	<.04	0.24	3.82	1, 208	.051
Extreme Aggression	0.06	5.73	1, 208	<.02	0.06	7.34	1, 208	.007
Controlling for Gender								
VDT: Time 1								
Unprovoked Aggression	0.28	12.92	1, 182	<.001	0.31	25.52	1, 182	<.001
Total Aggression	0.24	11.09	1, 182	.001	0.26	16.53	1, 182	<.001
Extreme Aggression	0.06	11.64	1, 182	.001	0.06	11.48	1, 182	.001
VDT: Time 2								
Unprovoked Aggression	0.20	3.15	1, 182	<.08	0.23	5.26	1, 182	.02
Total Aggression	0.11	0.94	1, 182	.33	0.14	2.85	1, 182	.09
Extreme Aggression	0.02	0.97	1, 182	.33	0.02	0.75	1, 182	.39

time task, all of which are commonly derived from the TAP (Anderson et al., 2004; Bartholow & Anderson, 2002; Bushman & Baumeister, 1998; Bushman, Ridge, Das, Key, & Busath, 2007; DeWall, Bushman, Giancola, & Webster, 2010; DeWall, Twenge, Bushman, Im, & Williams, 2010). First, we standardized and summed the intensity and duration settings from trial 1 to create a composite measure of *unprovoked aggression*. Second, we combined responses for intensity ($\alpha = .93$) and duration ($\alpha = .97$) levels participants selected across all of the trials to create a composite measure of *total aggression*. Third, we counted the number of times participants subjected their interaction partner to the extremely high noise intensity levels (i.e., 9 or 10) to create a measure of *extreme aggression*.

Procedure. At Time 1, participants came to the laboratory individually for a large study on interpersonal relationships and personality. One of the tasks they completed at this session was the VDT. As in Study 4, participants returned to the laboratory 4 weeks later. They returned with a close friend, performed the VDT again, and then engaged in the competitive reaction-time task ostensibly against their study partner.

Results and Discussion

First, we sought to replicate the Study 3 effect demonstrating strong test-retest reliability of pin insertion in the VDT. As expected, pin insertion at time 1 correlated strongly with responses at time 2, $r = .50$,

$p < .001^3$. This strong correlation again suggests that the VDT measures relatively stable aggressive tendencies over time.

Next, to test our primary hypothesis, we conducted six Poisson regression analyses with pin insertion as the dependent measure, crossing the two time points with the three aggression measures (unprovoked, total aggression, and extreme aggression). As expected, unprovoked aggression, total aggression, and extreme aggression exhibited significant associations with pin insertion at both time points. Six additional analyses using negative binomial regression demonstrated similar results (Table III). Associations remained significant or marginally significant after controlling for participant gender using both regression approaches, with the exception of associations between Time 2 total and extreme aggression.

These findings complement the Study 3 laboratory-based evidence linking pin insertion in the VDT to *psychological* aggression with laboratory-based evidence linking pin insertion in the VDT to *physical* aggression. Participants with stronger tendencies toward physical aggression—unprovoked aggression, total

³To ensure that this correlation was not affected by the non-normal nature of our dependent variable, we again repeated this analysis using both Poisson and negative binomial regression. Both regressions yielded nearly identical results [Poisson: $\chi^2(1, 208) = 99.39$, $p < .001$; Negative Binomial: $\chi^2(1, 208) = 131.58$, $p < .001$].

aggression, and extreme aggression—in the TAP inserted more pins into a voodoo doll representing their close friend.

STUDY 5: ASSOCIATIONS WITH LABORATORY-BASED PSYCHOLOGICAL AND PHYSICAL AGGRESSION

Study 5 employed quite different laboratory procedures to replicate the effects of Studies 3 and 4. Participants first completed the VDT with their relationship partner (romantic partner or close friend) represented by the doll. Next, they engaged in a videotaped conflict discussion task with their romantic partner or close friend. Trained coders rated these conflict discussions for how much anger participants conveyed during the task (psychological aggression), and they provided their best estimate of how likely it is that participants generally tend to be physically aggressive toward their partner (physical aggression). We predicted that participants who exhibited more anger toward their partner, and who conveyed the sense that they are physically aggressive toward their partner, during this task would insert a larger number of pins in the VDT. Though the design of this study does not permit us to make causal inferences, it will enable us to determine whether VDT responses relate to actual psychological and physical aggressive responses during a conflict discussion task.

Method

Participants. Eighty-nine undergraduates (56 female) participated in the study for partial course credit in conjunction with their relationship partner. Average age was 20.31 ($SD = 2.29$). Participants reported on their relationship with a romantic partner or close friend. For these individuals relationship length was: 40.4% 2+ years, 8.0% 19–24 months, 5.7% 13–18 months, 24.1% 7–12 months, 6.9% 4–6 months, 1.1% 2–3 months, and 12.6% less than 2 months. In addition, results did not differ across relationship type. Responses were therefore collapsed across relationship type.

Measure. Participants completed the VDT using the same instructions as in the previous studies. The average number of pins inserted into the doll was 5.42 ($SD = 10.22$). Overall, 44.9% of participants did not insert any pins into the voodoo doll, 41.8% inserted 1–10 pins, and 13.3% inserted more than 10 pins.

Procedure. Participants came to the laboratory with their romantic relationship partner or close friend, although the experimenter immediately led them to separate, individual cubicles. Participants then completed the VDT about their relationship partner. Afterward, the experimenter gave participants a sheet of paper and instructed both participants and their partners to write

about two or three of the most irritating issues that had recently occurred in their relationship with their partner. After five minutes of writing about these relationship issues, participants and partners entered a room that was equipped with a video camera. The experimenter instructed the participant and his or her partner that they would complete a videotaped, two-part discussion task about their relationship. In the first part (“argument phase”), participants and their partners took turns discussing the irritating issues in their relationship for 5 min.

After completing the first portion of the task, participants and their partners were placed into separate rooms where they completed tasks unrelated to the current investigation. The second part (“reconciliation phase”) involved participants and their partners returning to the video room and attempting to resolve their relationship issues over the next 4 min.

Coder ratings. Five trained coders, who were blind to the study hypotheses and participants’ behavior during the VDT, watched the argument phase and rated how much anger participants exhibited toward their partner (“How angry was the person during the initial argument phase?” $ICC = .93$). Coders then watched the reconciliation part of the discussion and rated how much anger participants exhibited (“How angry was the person during the reconciliation portion of the video?” $ICC = .92$). Finally, coders performed a more challenging task: They provided their best estimate of how likely it is that participants generally tend to be physically aggressive toward their partner (“How likely do you think it is that this person is currently being physically aggressive with their partner?” $ICC = .50$).

Results and Discussion

To test our hypothesis, we conducted three Poisson regression analyses using participants’ pin insertion as the dependent measure, two from coder ratings of participants’ anger (psychological aggression) and the other from coder estimates of participants’ general tendency to be physically aggressive toward their partner (physical aggression). As expected, higher levels of anger during both the argument phase, the reconciliation phase, and higher levels of estimated physical aggressiveness, were significantly associated with pin insertion in the VDT. These effects remained significant after controlling for relationship type ($ps < .05$). All effects remained significant after controlling for gender. Additional analyses using negative binomial regression yielded similar results (Table IV).

These findings replicate, with new procedures, both the Study 3 laboratory-based findings regarding psychological aggression (name-calling behavior) and the Study 4 laboratory-based findings regarding physical aggression (noise blast behavior). Study 5 participants who

TABLE IV. Voodoo Doll Task Responses Correlate With Coder-Rated Anger and Physical Aggressiveness (Study 5) and Major and Minor Intimate Partner Violence Perpetration Over the Past Year and Past 5 Years (Study 6)

Variable	<i>B</i>	χ^2	<i>df</i>	<i>p</i> -Value	<i>B</i>	χ^2	<i>df</i>	<i>p</i> -Value
Study 5	Poisson				Negative binomial			
Anger: Argument	0.40	5.25	1, 87	<.05	0.34	4.77	1, 87	<.05
Anger: Reconciliat.	0.38	7.37	1, 84	<.01	0.43	9.16	1, 84	<.01
Physical Aggression	0.83	9.86	1, 84	<.01	0.95	6.20	1, 84	.01
Controlling for Gender								
Anger: Argument	0.41	5.82	1, 81	<.05	0.35	4.51	1, 81	<.05
Anger: Reconciliat.	0.38	8.42	1, 81	<.01	0.41	9.19	1, 81	<.01
Physical Aggression	0.80	8.44	1, 81	<.01	0.89	5.73	1, 81	<.05
Study 6								
Maj. Phys. Assault: 1 year	0.04	14.64	1, 525	<.001	0.03	13.96	1, 525	<.001
Min. Phys. Assault: 1 year	0.06	22.96	1, 525	<.001	0.06	33.66	1, 525	<.001
Maj. Psych. Aggress.: 1 year	0.06	20.38	1, 525	<.001	0.07	34.77	1, 525	<.001
Min. Psych. Aggress.: 1 year	0.14	31.60	1, 525	<.001	0.13	21.29	1, 525	<.001
Maj. Phys. Assault: 5 years	0.04	21.10	1, 525	<.001	0.04	27.06	1, 525	<.001
Min. Phys. Assault: 5 years	0.06	27.67	1, 525	<.001	0.07	41.34	1, 525	<.001
Maj. Psych. Aggress.: 5 years	0.07	40.38	1, 525	<.001	0.09	52.25	1, 525	<.001
Min. Psych. Aggress.: 5 years	0.12	25.77	1, 525	<.001	0.11	19.87	1, 525	<.001

exhibited more anger (psychological aggression) toward their partner during both the argument and the reconciliation discussions, and who conveyed hints of a general tendency to be violent toward their partner (physical aggression), inserted more pins into a voodoo doll representing either their romantic partner or their close friend.

STUDY 6: ASSOCIATIONS WITH SEVERE AND MINOR PSYCHOLOGICAL AND PHYSICAL AGGRESSION

Study 6 had three goals. First, it sought to provide additional evidence for the construct validity of the VDT by replicating and extending our previous results. Participants completed the revised Conflict Tactics Scales (CTS2; Straus et al., 1996), which included measures of major physical assault and psychological aggression (not included in Study 2), as well as the measures for minor physical assault and psychological aggression.

Second, Study 6 sought to demonstrate that VDT responses would relate to physical assault and psychological aggression over a longer time period than was included in our previous studies. We predicted VDT responses from physical assault and psychological aggression committed over the past year and the past 5 years. Third, Study 6 sought additional evidence for the external validity of the VDT. Participants were a large community sample of individuals who had been in a committed relationship for at least 5 years. They were recruited from a national online survey administration service. We expected to replicate the results from the college student samples used in Studies 1–5.

Method

Participants. Five hundred twenty-five individuals (274 women, 251 men) participated in this study. To participate in the study, participants had to report involvement in a romantic relationship lasting at least five years. Average age was 43.27 ($SD = 31.02$). For these individuals, relationship length was: 38.2% 5–7 years, 12% 7.1–10 years, 8.6% 10.1–13 years, 5.5% 13.1–15 years, 7.6% 15.1–20 years, and 27.9% more than 20 years. Participants were recruited from the national online survey administration service, Zoomerang (www.Zoomerang.com), and they completed the self-report and voodoo task measures online, reporting on their romantic partners.

Measures

VDT. Participants completed the online version of the VDT used in Study 2. The average number of pins inserted into the doll was 2.80 ($SD = 8.39$). In all, 72% of participants did not insert any pins, 20.8% inserted 1–10 pins, and 7.2% inserted more than 10 pins.

Convergent Validity Measures. Participants completed Straus and colleagues' (1996) measures of *major physical assault* (e.g., "Choked my partner"), *minor physical assault* (e.g., "Slapped my partner"), *major psychological aggression* (e.g., "Called my partner fat or ugly"), and *minor psychological aggression* (e.g., "Insulted or swore at my partner"). These measures assessed how often participants had enacted physically or psychologically aggressive behaviors toward their partner over the past year and past five years. Responses were summed to create composite major physical assault, minor physical assault, major psychological aggression, and minor psychological aggression indices.

TABLE V. Voodoo Doll Task Responses Correlate With Major and Minor Intimate Partner Violence Perpetration Over the Past Year and Past 5 Years (Study 6) and Partner Provocation (Study 7)

Variable	<i>B</i>	χ^2	<i>df</i>	<i>p</i> -Value	<i>B</i>	χ^2	<i>df</i>	<i>p</i> -Value
Study 6								
Controlling for Gender								
Maj. Phys. Assault: 1 year	0.04	15.57	1, 524	<.001	0.04	15.32	1, 524	<.001
Min. Phys. Assault: 1 year	0.06	21.96	1, 524	<.001	0.07	29.10	1, 524	<.001
Maj. Psych. Assault: 1 year	0.06	19.61	1, 524	<.001	0.07	27.00	1, 524	<.001
Min. Psych. Assault: 1 year	0.15	32.94	1, 524	<.001	0.14	26.67	1, 524	<.001
Maj. Phys. Assault: 5 years	0.05	21.11	1, 524	<.001	0.05	26.56	1, 524	<.001
Min. Phys. Assault: 5 years	0.07	27.77	1, 524	<.001	0.08	35.98	1, 524	<.001
Maj. Psych. Assault: 5 years	0.07	37.39	1, 524	<.001	0.09	41.62	1, 524	<.001
Min. Psych. Assault: 5 years	0.12	27.94	1, 524	<.001	0.12	24.62	1, 524	<.001
Study 7								
Controlling for Gender								
State Provocation	0.40	7.93	1, 50	<.001	0.50	6.14	1, 50	<.001
State Provocation	0.40	7.94	1, 49	<.001	0.50	6.16	1, 49	<.001
Controlling for Trait Agg.								
State Provocation	0.40	7.92	1, 49	<.001	0.50	6.11	1, 49	<.001

Procedure. Participants completed the major and minor physical assault and psychological aggression scales and the VDT online through Zoomerang.

Results and Discussion

We predicted that VDT responses would relate to both major and minor physical assault and psychological aggression over the past year and the past five years. To test our hypotheses, we conducted eight Poisson regression analyses that used pin insertion as the dependent measure, two including major physical assault as the predictor (1 and 5 years), two including minor physical assault as the predictor (1 and 5 years), two including major psychological aggression as the predictor (1-year and 5-years), and two including minor psychological aggression as the predictor (1-year and 5-years). As expected, all predictors were significantly associated with stabbing the doll with more pins. These effects remained significant after controlling for gender. Additional analyses using negative binomial regression showed similar results (Tables IV and V).

In summary, these results demonstrate additional convergent and external validity of the VDT. With a large community sample of adults, self-report measures of major physical assault, minor physical assault, major psychological aggression, and minor psychological aggression toward a romantic partner over the past year and past 5 years correlated with responses on the VDT. Participants with stronger tendencies toward physical assault and psychological aggression inserted more pins into a voodoo doll representing their partner. These findings suggest that VDT responses relate to violent behavior that can cause severe physical injury.

One possible alternative explanation to this study is that VDT responses biased memory of prior IPV

behavior. That is, by inserting more pins into the doll, participants may have overestimated the number of times they perpetrated physical and psychological aggression against their partner. Although we believe it is unlikely that the VDT could exert such a strong bias on memory of prior IPV perpetration, Studies 7 and 8 aimed to show that the VDT could be used to assess state-level aggressive tendencies. If the VDT merely distorts perceptions and memories in a manner that is linked to higher aggression, then VDT responses should not shift as a function of factors that often increase aggression.

STUDY 7: PIN INSERTION FROM PROVOCATION AND DISPOSITIONAL AGGRESSIVENESS

The preceding studies have demonstrated associations of pin insertion in the VDT with physically and psychologically aggressive behavior, as assessed with both validated self-report measures (Studies 1, 2, and 6) and laboratory-based behavioral measures (Studies 3–5). The theoretical rationale underlying all of these studies was that pin insertion taps general, context-irrelevant tendencies toward physical and psychological aggression, at least regarding the person represented by the voodoo doll (e.g., one's romantic partner).

In Studies 7 and 8, we pursued two new goals.⁴ First, we sought to demonstrate that the VDT is also a sensitive and compelling measure of state-level aggressive tendencies. Toward this goal, we correlated pin insertion with provoking behavior from the romantic partner (the voodoo doll represented the romantic partner in these two

⁴ Variables from the Study 7 data set were included as part of a three-way interaction published in a previous report (Finkel et al., 2012). The two-way interaction reported below was not reported previously.

studies), as being provoked is among the strongest and most reliable predictors of aggression (Anderson & Bushman, 2002; Berkowitz, 1993). If participants who have (vs. have not) been provoked insert a greater number of pins into the doll representing their partner, that will suggest that scholars can use the VDT to assess state-level aggressive inclinations.

Our second goal for Studies 7 and 8 was to demonstrate that state-level factors continue to predict VDT responses after controlling for trait-level factors. In Study 7, we tested this possibility by examining whether provocation would relate to stabbing the voodoo doll with more pins, even after controlling for trait physical aggression.

Method

Participants. Fifty-one undergraduate heterosexual dating couples (102 individuals) volunteered to take part in the current study for monetary compensation (\$150/couple). Participants were 18.76 years old ($SD = 1.17$), and had been dating their current partner for 20.55 months ($SD = 17.80$), on average. One couple broke up before the end of the study, so the final dataset included 50 couples (100 individuals).

Measures

VDT. Participants were shown online pictures of a voodoo doll that represented their romantic partner. The number of pins stuck in the voodoo doll varied across each picture from 0 to 51. Participants selected the picture that represented the number of pins that they wished to stick into the voodoo doll. The average number of pins inserted into the doll was 1.89 ($SD = 6.68$).

Daily Provocation. On each daily questionnaire, they completed a face valid, one-item measure assessing the degree to which their partner made them feel “provoked” over the preceding 24-hr period ($-4 = \textit{Far less than usual}$, $0 = \textit{Typical for me}$, $+4 = \textit{Far more than usual}$; $M = -0.87$, $SD = 1.60$).

Trait Physical Aggression. At study intake, participants completed the Buss and Perry (1992) measure of *Physical Aggressiveness* ($M = 3.04$, $SD = 1.26$; $\alpha = .87$).

Procedure. This study was part of a larger, Internet-based investigation of romantic relationships among college students. After completing intake procedures, which included the Aggression Questionnaire (Buss & Perry, 1992), participants completed daily diaries online for 35 days (5 weeks); these diaries included the daily measure of partner provocation and the daily VDT. On average, participants completed 12.06 daily diaries ($SD = 8.05$).

Results and Discussion

We first sought to establish the reliability of the daily VDT across the 35 days of the study. Because our data

had a nested structure (i.e., days nested with people, people nested with couples), we used multilevel modeling to analyze our data (Raudenbush, Bryk, Cheong, & Congdon, 2000). The reliability of the VDT across the 5-week period, as estimated by HLM 6.08 (Raudenbush et al., 2000), was .75. This finding suggests that participants were consistent across time in their responses to the daily VDT.

Next, to test our hypothesis that the VDT is a sensitive measure of state-level aggressive tendencies, we conducted a Poisson regression analysis predicting the number of pins participants inserted in the voodoo doll as the outcome measure from participants’ of the degree to which their partner had provoked them that day ($M = 0$, $SD = 1$). All analyses reported below used group-mean centering.

As predicted, pin insertion on the VDT was reliably associated with state provocation, such that participants stuck more pins into the voodoo doll when they felt provoked by their partner compared to when they felt less provoked. This effect remained significant after controlling for participant gender and trait physical aggression. Consistent with Study 1, participants who reported higher levels of aggressiveness stuck a greater number of pins in the voodoo doll than their less aggressive counterparts, $B = 0.64$, $t(49) = 2.40$, $p = .01$. Negative binomial regression analyses yielded similar results, including the association between trait physical aggression and pin insertion, $B = 0.60$, $t(49) = 2.31$, $p = .02$, again emerged as significant predictors of VDT pin usage (Table V).

In sum, this daily diary study revealed three central findings. First, it showed that pin insertion on the VDT is reliable from one day to the next over a 5-week period. Second, it replicated the Study 1 association of dispositional physical aggressiveness with pin insertion. These first two findings lend additional support for our assertion that the VDT functions well as a measure of stable aggressive tendencies over time. Third, this study provided the first demonstration that the VDT also functions well as a state-level measure of aggression.

STUDY 8: USING EXPERIMENTAL PROCEDURES TO ASSOCIATE PIN INSERTION FROM

Provocation and Dispositional Aggressiveness

Study 8 sought to replicate the key results from Study 7, this time experimentally manipulating partner provocation rather than assessing it with participants’ self-reports. Participants first completed the Buss and Perry (1992) measure of dispositional physical aggressiveness. Next, adapting procedures from Finkel et al. (2009), we asked participants to complete a drawing task and ostensibly received either insulting or neutral feedback

from a romantic partner before completing the VDT. Whereas Finkel et al. (2009) manipulated (dis)inhibition and instigation, the current study measured an impellor (i.e., dispositional physical aggressiveness) and manipulated an instigator (i.e., provocation).

To show how the VDT can be used in a setting with no computer, Study 8 participants completed an “in-person” version of the VDT rather than the online version used in the previous studies. That is, participants were given an actual voodoo doll to stab with real pins instead of completing the VDT using a computer image of a voodoo doll and virtual pins. Although we had no a priori reasons to expect differences between the two versions of the task, obtaining similar responses would provide convergent validity in support of the VDT as a measure of aggressive inclinations.

Method

Participants. Forty-three heterosexual undergraduate dating couples (86 individuals) volunteered to take part in the current study for monetary compensation (\$20/couple). Participants were 19.10 years old ($SD = 1.00$), and had been dating their current partner for 13.51 months ($SD = 25.38$), on average. Both members of the couple attended the laboratory session, although they completed all portions of the study separately.

Measures

VDT. The experimenter gave each participant a package of 200 straight pins and a voodoo doll that represented their romantic partner. She asked them to take 5 min to inflict harm on the voodoo doll as a way to get out any negative feelings they might be having before they left the laboratory. The average number of pins inserted into the doll was 10.22 ($SD = 14.51$). Overall, 41.4% of participants did not insert any pins, 25.3% inserted 1–10 pins, 16.1% inserted 11–25 pins, and 17.2% inserted more than 20 pins. In terms of the percentage of pins inserted, this is 5.1% of the total pins available, which is slightly higher than the average percentage of total pins (4.7%) inserted in the versions of the task used in Studies 1–7.

Trait physical aggressiveness. Participants completed the Buss and Perry (1992) measure of physical aggressiveness ($M = 2.80$, $SD = 1.24$; $\alpha = .86$).

Procedure. Participants completed all measures in a single laboratory session. After completing the physical aggressiveness measure, they performed a creativity task and were led to believe that their romantic partner was going to evaluate their creative abilities (see Finkel et al., 2009; Slotter et al., 2012). The experimenter gave them paper and colored pencils and asked them to draw the most creative picture that they could that included the following five objects: a house, a car, a tree, and two

people. After they completed the drawing, the experimenter collected it and told participants that their romantic partner, who was in a different room, would evaluate its creativity. She also told them that they would receive a monetary reward for more creative drawings, based on their partner’s evaluation. Participants could receive a maximum of \$5.00; however, that was only if their partner gave them the maximum creativity rating (1 = *not at all creative*; 10 = *extremely creative*; each scale point was worth US \$0.50 toward the participants’ monetary reward). Participants were told that they were going to complete a filler task while their partner evaluated their drawing, and that they would have the chance to evaluate their partner’s drawing later on in the experiment.

In reality, participants were randomly assigned to receive false feedback—either provoking or non-provoking—about their creative drawing to manipulate their experience of partner provocation. In the *provoking feedback* condition, participants received feedback that their partner had rated their drawing as a three on the creativity scale, which corresponded to \$1.50. Additionally, the experimenter told the participants in this condition that their partner had commented that their drawing was “not that great and really boring,” and that they were “disappointed in [the participant].” In contrast, participants in the *non-provoking feedback* condition learned that their partner had rated their drawing as a 9 on the creativity scale, which corresponded to \$4.50. Additionally, the experimenter told the participants in this condition that their partner had commented that their drawing was “fantastic and really creative,” and said that they were “really proud of [the participant].”

After receiving the false feedback from their romantic partner, the experimenter told participants that because the feedback they received might have been upsetting to them, the researchers were required to give them a task that would allow them to release any negative energy they experienced during the study. The experimenter then left participants alone for 5 min with the voodoo doll representing their romantic partner (described above) and the package of straight pins. She instructed them to use the pins to do harm to the voodoo doll as a way of getting out any negative feelings they might be having as a result of the study before they left the laboratory, and that their behavior on this task was completely confidential. After 5 min, the experimenter dismissed participants and recorded the number of pins that participants had placed in the voodoo doll.

Results

To test our hypothesis that the VDT is a sensitive measure of aggressive inclinations, we conducted a Poisson regression analysis using the number of pins

participants inserted in the voodoo doll as the dependent measure from whether participants had received provoking or non-provoking feedback from their romantic partner (non-provoking feedback = -0.5 , provoking feedback = 0.5). We used multilevel modeling to account for the nesting of person within couple.

As predicted, provocation caused participants to stab the voodoo doll with more pins compared to when they were not provoked (Table VI). This effect remained significant after controlling for participant gender and trait physical aggression. As in Studies 1 and 7, trait physical aggressiveness also emerged as a significant predictor of VDT pin insertion such that participants who reported higher levels of aggressiveness stuck a greater number of pins in the voodoo doll than their less aggressive counterparts, $B = .20$, $t(42) = 3.76$, $p < .01$. Unlike our previous studies, additional analyses using negative binomial regression analysis revealed neither a significant main effect of provocation nor a significant association between VDT responses and trait physical aggression, $B = 0.22$, $t(42) = 1.00$, $p = .32$.

In sum, Study 8 provided some additional, causal evidence regarding the validity of the VDT as a state-level measure of aggressive inclinations. When participants experienced partner provocation, they stabbed the voodoo doll with more pins than participants who had not been provoked. This relationship emerged using Poisson regression but did not occur when using negative binomial regression. In general, the results from Study 8 offered additional evidence that a momentary experience of provocation by their romantic partner caused participants to stab a voodoo doll that represented their partner. It provided additional evidence that VDT responses correlate with a well-validated measure of individuals' physically aggressive tendencies.

STUDY 9: RULING OUT THE CATHARSIS FRAMING AS AN ALTERNATIVE EXPLANATION

The first eight studies offered consistent evidence that the VDT provides a reliable and valid measure of aggressive inclinations across multiple settings and relationship contexts. Study 9 sought to test the alternative explanation that these effects were due to catharsis framing from the task instructions to release negative affect instead of to an actual aggressive inclination. We examined whether the catharsis instructions would moderate the effect of trait physical aggression on pin insertion by experimentally manipulating the instructions that participants read before completing the VDT. Because we hypothesized that the VDT assesses aggressive inclinations rather than instructions-inspired catharsis motivation, we predicted VDT responses would not differ according to which instructions participants received.

Method

Participants. Ninety-eight individuals (51 women, 47 men) participated in this study. Average age was 30.07 ($SD = 9.17$). Participants were recruited from Amazon's Mechanical Turk website (www.mturk.com). Participants completed the study in exchange for \$0.10. Though this payment is relatively small, it is not unusual to pay Amazon Mechanical Turk participants this amount (Buhrmester, Kwang, & Gosling, 2011).

Measures

VDT. Participants completed the online version of the VDT used in the previous studies. The average number of pins inserted into the doll was 12.55 ($SD = 19.11$). In all, 49.5% of participants did not insert any pins into the doll, 23.1% inserted 1–10 pins, and 27.4% inserted more than 10 pins.

TABLE VI. Experimental Provocation Increases Voodoo Doll Task Responses (Study 8)

Variable	<i>B</i>	χ^2	<i>df</i>	<i>p</i> -Value	<i>B</i>	χ^2	<i>df</i>	<i>p</i> -Value
Study 8								
Controlling for Gender								
Experimental Provocation	0.61	8.19	1, 43	<.001	0.46	1.08	1, 43	.28
Experimental Provocation	0.86	6.59	1, 42	<.001	0.55	0.91	1, 42	.37
Controlling for Trait Agg.								
Experimental Provocation	0.48	5.78	1, 42	<.001	0.38	0.87	1, 42	.39
Study 9								
Trait Physical Aggression	0.74	47.95	1, 98	<.001	1.16	73.32	1, 98	<.001
Trait Physical Aggression × Catharsis Condition	-0.30	1.12	1, 98	.29	-0.63	5.35	1, 98	.02
Controlling for Gender								
Trait Physical Aggression	0.75	36.57	1, 97	<.001	0.97	54.67	1, 97	<.001
Trait Physical Aggression × Catharsis Condition	-0.29	1.39	1, 97	.24	-0.38	2.19	1, 97	.14

Voodoo doll task responses correlate with trait physical aggression and are not reliably moderated by catharsis instructions (Study 9).

Convergent Validity Measure. Participants completed the Buss and Perry (1992) measure of *Physical Aggressiveness*. Responses were averaged across items to form a composite of trait physical aggression ($M = 3.28$, $SD = 1.34$; $\alpha = .88$).

Procedure. Participants completed all aspects of the study online. After completing the measure of physical aggressiveness, participants were randomly assigned to receive either *cathartic* or *neutral* instructions prior to the VDT. The cathartic instructions were identical to those used in Studies 1–8. However, those in the *neutral* condition were given the following instructions:

Below, you'll be shown a doll that represents your current (or most recent) romantic partner. You will get to choose how many needles (up to 51) you would like us to put in the doll that represents your current or most recent romantic relationship partner.

After completing the VDT, participants were debriefed and compensated.

Results and Discussion

We predicted that trait physical aggressiveness would be associated with greater pin insertion on the VDT, and that this effect would not be moderated by the presence of cathartic instructions. To test our hypotheses, we conducted a Poisson regression analysis predicting pin insertion from trait physical aggressiveness, instruction condition ($0.5 = \textit{catharsis}$ condition, $-0.5 = \textit{neutral}$ condition), and an interaction between trait physical aggression and instruction condition. Simple effect tests were conducted using the pooled error variance term.

Replicating previous findings, trait physical aggressiveness was positively associated with pin insertion. Of primary relevance to the present study, this association was not moderated by instruction condition. Indeed, trait physical aggression was positively associated with pin insertion in both the catharsis condition, $B = 0.62$, $t(94) = 3.56$, $p < .001$, and the neutral condition, $B = 0.85$, $t(94) = 7.06$, $p < .001$. All effects were unchanged after controlling for gender (see Table VI).

Negative binomial regression analyses yielded largely similar results. Trait physical aggressiveness was positively associated with pin insertion. However, this association was moderated by instruction condition. Trait physical aggression was positively associated with pin insertion in both the catharsis condition, $B = 0.85$, $t(94) = 3.99$, $p < .001$, and the neutral condition, $B = 1.48$, $t(94) = 8.75$, $p < .001$. Yet, the interaction indicates that the association between trait physical aggression and pin insertion was much stronger for those in the neutral condition.

Thus, our findings suggest that the catharsis instructions, if anything, reduced the number of pins participants inserted. Analyses controlling for partner gender left

these effects unchanged, with the exception of rendering the interaction between trait physical aggression and catharsis instruction condition nonsignificant (see Table II). Trait physical aggression continued to have a positive association with pin insertion in both the catharsis condition, $B = 0.78$, $t(94) = 3.89$, $p < .001$, and the neutral condition, $B = 1.16$, $t(94) = 6.97$, $p < .001$. Together, these findings suggest that the VDT assesses aggressive inclinations rather than some sort of instructions-inspired catharsis motivation. Participants with stronger tendencies toward physical aggression inserted more pins into a voodoo doll representing their romantic partner regardless of whether they were told that pin insertion would get out “negative energy” or not.

GENERAL DISCUSSION

People behave aggressively for many reasons and toward many different people. To understand the root causes of aggression, researchers have developed various methods to assess aggressive inclinations. How aggression researchers do their work has evolved to include three different research “worlds” (i.e., no computer, computer in a laboratory, Internet-based research) and three different relationship contexts (i.e., strangers, nonromantic members of one’s social network, romantic partners). Despite the validity of many aggression measures, no measure to date has accommodated these changes in aggression research. The current article introduces a new method—the VDT, or *VDT*—that offers a measure of aggressive inclinations that has comparable validity across settings and relationship contexts.

Using research on magical beliefs and the law of similarity as a conceptual framework (Rozin et al., 1986), we propose that people transfer characteristics of a person onto a voodoo doll representing that person. By intentionally causing harm to the doll by stabbing it with pins, people may have a psychologically similar experience to actually causing harm to the person that the doll represents. If so, VDT responses should show high levels of reliability across time, and VDT responses should relate to a variety of constructs indicative of both trait and state aggressive tendencies.

Nine methodologically diverse studies established the VDT’s reliability and validity as a measure of aggressive inclinations across research settings and relationship contexts. For example, the VDT showed strong reliability. VDT responses showed high test–retest reliability over 4 weeks and within-person reliability on a daily level. In addition, the VDT showed substantial construct and convergent validity. The number of pins participants stuck into the doll was associated with several self-report indicators of aggression, including

measures of trait physical aggression, minor and major physical assault and psychological aggression perpetrated in a close relationship. Pin insertion was also associated with insulting a close relationship partner during a problem-solving task, expressing greater anger and showing higher aggressive tendencies during a conflict discussion task, and blasting a close relationship partner with louder and more prolonged noise during a competitive reaction-time task. Responses were quite consistent across online and in-person versions of the VDT, which underscore the robustness of the VDT as a measure of aggressive inclinations that researchers can use in various settings.

A situational factor known to increase aggressive behavior also increased the number of pins that participants inserted into the voodoo doll, suggesting that researchers can also use the VDT as a measure of state aggressive inclinations. Provocation, which is one of the most important causes of aggression (Anderson & Bushman, 2002; Berkowitz, 1993), increased pin insertion on the VDT. Participants who experienced daily provocation from a romantic partner inserted more pins into the voodoo doll that represented their partner. An experimental manipulation of provocation also increased pin insertion on the VDT.

Thus, the VDT addresses limitations of research setting and relationship type inherent in existing measures of aggressive inclinations. The VDT does not undermine the validity or reliability of these previous measures. Instead, the VDT offers researchers a new method to assess aggressive inclinations across the three major research settings and relationship types with which aggression research often takes place.

More broadly, the current work underscores the role of magical beliefs in understanding aggressive inclinations. Consistent with the law of similarity, when confronted with a doll that represented another person, participants consistently behaved in ways suggesting that the doll took on the actual qualities of the other person. Although magical beliefs play no role in major aggression theories (to our knowledge), the current findings suggest that much can be gained from appreciating the role of magical beliefs in informing our understanding of the causes of aggression.

Limitations and Future Directions

The present studies provided consistent evidence supporting the reliability and validity of the VDT. Yet there are some limitations to the VDT that warrant consideration. First, the current findings did not assess the underlying motivations associated with aggression. People behave aggressively for a variety of reasons, including their reactive responses under conditions of negative affect and instrumental responses in order to

obtain some benefit (Bushman & Anderson, 2001). Our findings suggest that the VDT can be used to assess inclinations toward reactive aggression, such as in response to provocation. Future research may examine whether the VDT can be used to assess aggressive inclinations that occur as a result of instrumental motives. For example, if the VDT were framed as a means of gaining a desired goal (e.g., money, recognition, revenge), we predict that the VDT could also be used to assess instrumental aggressive inclinations.

Another limitation is that the VDT may not provide a pure assessment of insertion intensity and location. A more intense response is coded as a higher number of pins inserted into the voodoo doll. Participants do not have the opportunity to adjust how forcefully they insert the pins. It is possible that participants predisposed to behave aggressively may strongly thrust one or two pins into the voodoo doll instead of inserting a large number of pins, if given the option. It is also an open question as to whether participants predisposed to behave aggressively would choose different insertion locations (e.g., belly, genitals, head). The VDT may be altered to allow participants to control insertion intensity and insertion location. Nonetheless, our results suggest that using the number of pins inserted is an effective means of assessing a variety of trait- and state-level indicators of aggressive inclinations.

A third limitation is that VDT responses had a non-normal distribution in each study, with the majority of participants selecting zero pins to insert into a doll representing their partner. To account for the non-normal nature of VDT responses, we conducted our analyses using the two most common approaches for count data: Poisson regression and negative binomial regression. Given the high number of zeros, the VDT might not be sensitive enough to distinguish variability at the highly nonaggressive end of the continuum. That said, these findings also provide additional evidence regarding the validity of the VDT. If participants did not take the VDT seriously, then they would have little problem inserting large numbers of pins into the doll. Instead, VDT responses mimicked responses to other count-based measures of aggressive inclinations in which the majority of participants often report zero acts of perpetration, such as the Revised Conflict Tactics Scale (Straus et al., 1996) and the Safe Dates Physical Violence Scale (Foshee et al., 1996).

A fourth limitation is that VDT responses may represent a cognitive dissonance reduction strategy, in which people recognize the VDT as a measure of aggression and therefore try to be consistent with their reports of other measures of aggression. This possibility, which also would apply to other measures of aggression, is unlikely. For example, participants who inserted a large

number of pins into the voodoo doll would not report perpetrating more acts of intimate partner violence on the Conflict Tactics scale because those acts of IPV occurred in the past. It is difficult to disentangle dissonance reduction with a genuine proclivity toward aggression; probably both are at play. But there is more reason to expect that consistency across multiple measures of aggression represent an assessment of a person's aggressiveness instead of a person's strategy to maintain cognitive consistency. For example, it is unlikely that an extremely peaceful person would stab the voodoo doll representing their partner with many pins and, in an effort to maintain consistency with that behavior, engage in acts of physical assault toward a close relationship partner. It is far more plausible that a person who stabs the voodoo doll with many pins will behave aggressively toward their partner because the person has aggressive tendencies.

Another limitation is that the current findings may be compromised because of participant suspicion or demand characteristics. We limited suspicion by not deceiving participants about the nature of the task. We also limited demand characteristics by telling participants that they could insert as many or as few pins as they wished and by leaving the room while participants completed the study. In eight of the nine studies, participants completed the VDT on a computer, which further reduced any demand characteristics because the computer recorded their responses instead of the experimenter doing so.

Another limitation is that we observed some variability across the samples in the number of pins inserted. We speculate that two reasons may help explain this variability. First, participants in one study were given a larger number of possible pins to insert into the doll (i.e., 200 in Study 8) compared to the smaller number of possible pins (i.e., 51) in the other eight studies. Therefore, Study 8 participants may have exerted a larger number of pins than other participants because they perceived their responses as a function of the total percentage of pins they could insert. Second, the largest number of pins inserted and largest variability in responses was found in Study 9, which used participants from Amazon Mechanical Turk. Historically, Amazon Mechanical Turk participants are more diverse than typical online samples in terms of their age, nationality, and ethnicity (Buhrmester et al., 2011). Seven of our other studies used undergraduates from American universities, whereas Study 6 used a national American sample of adults. These samples were likely more homogenous than the Amazon Mechanical Turk sample. We did not anticipate such differences across samples and, hence, did not include measures that would enable us to directly examine these possible explanations. Future research may explore the reasons underlying these differences. Despite these differences in average number

of pins inserted across samples, there was consistent evidence across all samples that VDT responses could be used as a valid and reliable measure of aggressive inclinations.

A final limitation is that the current studies did not assess whether individual differences in superstitious (or other magical) thinking may moderate our effects. Although VDT responses may be especially pronounced among highly superstitious participants, prior work has argued that the law of similarity and magical beliefs characterize psychological processes in ordinary people instead of being confined exclusively to superstitious people (Pronin et al., 2006; Rozin et al., 1986). Indeed, the VDT's consistent effects across various methods and measures, and in the absence of a measure of superstition, offer additional support that a twinkle of magical beliefs exist among the typical well-educated person and therefore can be used to understand the causes of aggression in general samples.

Concluding remarks

To understand why people behave aggressively, researchers must use reliable and valid measures of aggressive inclinations. The current investigation introduced and validated a new behavioral method for assessing aggressive inclinations toward strangers and close relationship partners, which we call the *voodoo doll task*. The VDT represents an important step in the study of aggressive inclinations. By having a reliable, valid, and easy-to-use measure of aggressive inclinations that is applicable across research settings and relationship contexts, researchers will be better equipped to uncover when and why people have aggressive inclinations—and potentially how such aggression can be reduced.

ACKNOWLEDGMENTS

This research was supported in part by grants from the National Science Foundation (BCS-719780 to E.J.F. and BCS-1104118 to C.N.D.) and start-up funds from the University of Kentucky College of Arts and Sciences (C. M.R.). The opinions and conclusions expressed herein are those of the authors and do not necessarily reflect the opinions of the NSF.

REFERENCES

- Anderson, C. A., Buckley, K. E., & Carnagey, N. L. (2008). Creating your own hostile environment: A laboratory examination of trait aggression and the violence escalation cycle. *Personality and Social Psychology Bulletin*, *34*, 462–473.
- Anderson, C. A., & Bushman, B. J. (1997). External validity of “trivial” experiments: The case of laboratory aggression. *Review of General Psychology*, *1*, 19–41.
- Anderson, C. A., & Bushman, B. J. (2002). Human aggression. *Annual Review of Psychology*, *53*, 27–51.

- Anderson, C. A., Carnagey, N. L., Flanagan, M., Benjamin, A. J., Eubanks, J., & Valentine, J. C. (2004). Violent video games: Specific effects of violent content on aggressive thoughts and behavior. *Advances in Experimental Social Psychology*, 36, 199–249.
- Archer, J. (2000). Sex differences in aggression between heterosexual partners: A meta-analytic review. *Psychological Bulletin*, 126, 697–702.
- Babcock, J. C., Costa, D. M., Green, C. E., & Eckhardt, C. I. (2004). What situations induce intimate partner violence? A reliability and validity study of the Proximal Antecedents to Violent Episodes (PAVE) scale. *Journal of Family Psychology*, 18, 433–442.
- Baron, R. A., & Richardson, D. R. (1994). *Human aggression* (2nd ed.). New York: Plenum Press.
- Bartholow, B. D., & Anderson, C. A. (2002). Effects of violent video games on aggressive behavior: Potential sex differences. *Journal of Experimental Social Psychology*, 38, 283–290.
- Berkowitz, L. (1993). *Aggression: Its causes, consequences, and control*. New York: McGraw-Hill.
- Buhrmester, M. R., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality data? *Perspectives on Psychological Science*, 6, 3–5.
- Bushman, B. J., & Anderson, C. A. (2001). Is it time to pull the plug on the hostile versus instrumental aggression dichotomy? *Psychological Review*, 108, 273–279.
- Bushman, B. J., & Baumeister, R. F. (1998). Threatened egotism, narcissism, self-esteem, and direct and displaced aggression: Does self-love or self-hate lead to violence? *Journal of Personality and Social Psychology*, 75, 219–229.
- Bushman, B. J., & Huesmann, L. R. (2010). Aggression. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (5th ed., pp. 833–863). New York: John Wiley & Sons.
- Bushman, B. J., Ridge, R. D., Das, E., Key, C. W., & Busath, G. L. (2007). When God sanctions killing; Effect of scriptural violence on aggression. *Psychological Science*, 18, 204–207.
- Buss, A. H., & Perry, M. P. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology*, 63, 452–459.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.
- Damisch, L., Stoberock, B., & Mussweiler, T. (2010). Keep your fingers crossed! How superstition improves performance. *Psychological Science*, 21, 1014–1020.
- Denzler, M., Förster, J., & Liberman, N. (2009). How goal-fulfillment decreases aggression. *Journal of Experimental Social Psychology*, 45, 90–100.
- DeWall, C. N., Bushman, B. J., Giancola, P. R., & Webster, G. D. (2010). The big, the bad, and the boozed-up: Weight moderates the effect of alcohol on aggression. *Journal of Experimental Social Psychology*, 46, 619–623.
- DeWall, C. N., Twenge, J. M., Bushman, B. J., Im, C., & Williams, K. D. (2010). Acceptance by one differs from acceptance by none: Applying social impact theory to the rejection-aggression link. *Social Psychological and Personality Science*, 1, 168–174.
- Eckblad, M., & Chapman, L. J. (1983). Magical ideation as an indicator of schizotypy. *Journal of Counseling and Clinical Psychology*, 51, 215–225.
- Eckhardt, C. I., Barbour, K. A., & Davison, G. C. (1998). Articulated thoughts of maritally violence and nonviolent men during anger arousal. *Journal of Consulting and Clinical Psychology*, 66, 259–269.
- Finkel, E. J., DeWall, C. N., Slotter, E. B., McNulty, J. K., Pond, R. S., Jr., & Atkins, D. C. (2012). Using I³ theory to clarify when dispositional aggressiveness predicts intimate partner violence perpetration. *Journal of Personality and Social Psychology*, 102, 533–549.
- Finkel, E. J., DeWall, C. N., Slotter, E. B., Oaten, M., & Foshee, V. A. (2009). Self-regulatory failure and intimate partner violence perpetration. *Journal of Personality and Social Psychology*, 97, 483–499.
- Foshee, V. A., Linder, F. F., Bauman, K. E., Langwick, S. A., Arriaga, X., Heath, J., et al. (1996). The Safe Dates project: Theoretical basis, evaluation design, and selected baseline findings. *American Journal of Preventative Medicine*, 12, 39–48.
- Gardner, W., Mulvey, E. P., & Shaw, E. C. (1995). Regression analyses of counts and rates: Poisson, overdispersed poisson, and negative binomial models. *Psychological Bulletin*, 118, 392–404.
- Gendler, T. S. (2008). Alief in action (and reaction). *Mind & Language*, 23, 552–585.
- Giancola, P. R., & Chermack, S. T. (1998). Construct validity of laboratory aggression paradigms: A response to Tedeschi and Quigley (1996). *Aggression and Violent Behavior*, 3, 237–253.
- Giancola, P. R., & Parrott, D. J. (2008). Further evidence for the construct validity of laboratory aggression paradigms. *Aggressive Behavior*, 34, 214–229.
- King, L. A., Burton, C. B., Hicks, J. A., & Dragotis, S. (2007). Ghosts, UFOs, and magic: Positive affect and the experiential system. *Journal of Personality and Social Psychology*, 92, 905–919.
- Kramer, T., & Block, L. (2008). Conscious and nonconscious components of superstitious beliefs in judgment and decision making. *Journal of Consumer Research*, 34, 783–793.
- Lieberman, J. D., Solomon, S., Greenberg, J., & McGregor, H. A. (1999). A hot new way to measure aggression: Hot sauce allocation. *Aggressive Behavior*, 25, 331–348.
- Lindeman, M., & Aarnio, K. (2007). Superstitious, magical, and paranormal beliefs: An integrative model. *Journal of Research in Personality*, 41, 731–744.
- Long, J. S. (1997). *Regression models for categorical and limited dependent variables*. Thousand Oaks, CA: Sage Publications.
- MacDonald, K. B. (2008). Effortful control, explicit processing, and the regulation of human evolved predispositions. *Psychological Review*, 115, 1012–1031.
- McKay, R. T., & Dennett, D. C. (2009). The evolution of misbelief. *Behavioral and Brain Sciences*, 32, 493–561.
- Nemeroff, C., & Rozin, P. (2000). The makings of the magical mind. In K. S. Rosengren, C. N. Johnson, & P. L. Harris (Eds.), *Imagining the impossible: magical, scientific, and religious thinking in children* (pp. 1–34). New York: Cambridge University Press.
- Ng, T., Chong, T., & Du, X. (2010). The value of superstitions. *Journal of Economic Psychology*, 31, 293–309.
- Pronin, E., Wegner, D. M., McCarthy, K., & Rodriguez, S. (2006). Everyday magical powers: The role of apparent mental causation in the overestimation of personal influence. *Journal of Personality and Social Psychology*, 91, 218–231.
- Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., & Congdon, R. T. (2000). HLM (Version 6.8) [Software]. Lincolnwood, IL: Scientific Software International.
- Risen, J. L., & Gilovich, T. (2008). Why people are reluctant to tempt fate. *Journal of Personality and Social Psychology*, 95, 293–307.
- Rozin, P., Millman, L., & Nemeroff, C. (1986). Operation of the laws of sympathetic magic in disgust and other domains. *Journal of Personality and Social Psychology*, 50, 703–712.
- Schafer, J., Caetano, R., & Clark, C. L. (1998). Rates of intimate partner violence in the United States. *American Journal of Public Health*, 88, 1702–1704.
- Slotter, E. B., Finkel, E. J., DeWall, C. N., Lambert, N. M., Pond, R. S., Bodenhausen, G. V., & Fincham, F. D. (2012). Putting the brakes on aggression toward a romantic partner: The inhibitory influence of relationship commitment. *Journal of Personality and Social Psychology*, 102, 291–305.
- Sluka, J. (1992). The anthropology of conflict. In C. Nordstrom & J. Martin (Eds.), *The paths to domination, resistance, and terror* (pp. 18–36). Berkeley: University of California Press.
- Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, 8, 220–247.

- Strahan, R., & Gerbasi, K. C. (1972). Short, homogeneous versions of the Marlowe–Crowne social desirability scale. *Journal of Clinical Psychology, 28*, 191–193.
- Straus, M. A. (1979). Measuring intrafamily conflict and violence: The conflict tactics scales. *Journal of Marriage and the Family, 41*, 75–88.
- Straus, M. A. (1995). *Manual for the conflict tactics scales*. Durham, NH: Family Research Laboratory, University of New Hampshire.
- Straus, M. A., Hamby, S. L., Boney-McCoy, S., & Sugarman, D. B. (1996). The revised conflict tactics scales (CTS2): Development and preliminary psychometric data. *Journal of Family Issues, 17*, 283–316.
- Taylor, S. (1967). Aggressive behavior and physiological arousal as a function of provocation and the tendency to inhibit aggression. *Journal of Personality, 35*, 297–310.
- Woolley, J. D. (1997). Thinking about fantasy: Are children fundamentally different thinkers and believers from adults? *Child Development, 68*, 991–1011.
- Zeigarnik, B. (1927). Über das Behalten von erledigten und unerledigten Handlungen [On the retention of completed and uncompleted transactions]. *Psychologische Forschung, 9*, 1–85.
- Zusne, L., & Jones, W. H. (1989). *Anomalistic psychology: A study of magical thinking*. Hillsdale, NJ: Erlbaum.