

I³ Theory:

Instigating, Impelling, and Inhibiting Factors in Aggression

Erica B. Slotter and Eli J. Finkel

Northwestern University

Interpersonal aggression is prevalent and disturbing. This chapter presents a meta-theoretical perspective, *I³ Theory*, that seeks (a) to impose theoretical coherence on the massive number of established risk factors for aggression and (b) to employ the tools of statistical (and conceptual) moderation to gain new insights into the processes by which a previously nonaggressive interaction escalates into an aggressive one (see Finkel, 2007, 2008). *I³ Theory* (pronounced “I-Cubed Theory”) does not advance one key variable (or even a specific set of key variables) as the root cause of aggression. Rather, it seeks to present an organizational structure for understanding both (a) the process by which a given risk factor promotes aggression and (b) how these risk factors interrelate to aggravate or mitigate the aggression-promoting tendencies of each. As detailed below, *I³ Theory* suggests that scholars can predict whether an individual will behave aggressively in a given situation by examining the main and interactive effects of the Instigating triggers, Impelling forces, and Inhibiting forces at play.

Aggression refers to any behavior carried out with the primary proximal goal of inflicting physical harm on a target who is motivated to avoid being harmed (Baron & Richardson, 1994).

(We do not examine in this chapter other forms of aggression, such as verbal, relational, or sexual aggression.) In a recent literature review, C. A. Anderson and Bushman (2002, p. 29) report that “five main theories of aggression guide most current research”: (a) Cognitive Neoassociation Theory, (b) Excitation Transfer Theory, (c) Social Learning Theory, (d) Script Theory, and (e) Social Interaction Theory. In the ensuing paragraphs, we briefly review the core tenets of each of these five theories and of a recent integration called the General Aggression Model (C. A. Anderson & Bushman, 2002; DeWall & C. A. Anderson, Chapter 1, this volume). Thereafter, we introduce I³ Theory and then briefly and selectively review the aggression literature from its perspective.

Extant Theories of Aggression

According to *Cognitive Neoassociation Theory* (which largely subsumes the earlier Frustration-Aggression Hypothesis; Dollard, Doob, Miller, Mowrer, & Sears, 1939), negative affective states cause people to aggress (Berkowitz, 1993). Negative affect, which is frequently caused by an aversive event such as frustration, provocation, or physical pain, can automatically trigger cognitive, affective, physiological, and even motor tendencies associated with aggression. Cognitive Neoassociation Theory also emphasizes that higher-level cognitive processes, such as attributions and appraisals, can modulate these automatic “instigations to aggress.” According to *Excitation Transfer Theory* (Zillman, 1983), misattributed physiological arousal leads to aggression. If individuals are experiencing incidental (unrelated) physiological arousal when they encounter an anger-inducing stimulus, their anger may become exacerbated by the misattribution of the incidental arousal to this stimulus. This exacerbated anger can, in turn, precipitate aggression.

According to *Social Learning Theory* (Bandura, 1973), individuals develop aggressive tendencies when they observe other individuals behave aggressively, particularly if those other individuals are likable, have high status, are rewarded for their aggressive behavior, and so forth. According to *Script Theory* (Huesmann, 1998; Huesmann, Dubow, & Boxer, Chapter 8, this volume), individuals learn aggressive scripts from the world around them, including from the mass media. Scripts, which refer to well-learned, highly accessible, and integrated mental representations of situation-specific causal links, goals, and action plans (Schank & Abelson, 1977), can cause individuals to engage in complex aggressive acts once circumstances trigger a key aspect of the aggressive script. Finally, according to *Social Interaction Theory* (Tedeschi & Felson, 1994), individuals enact aggressive behavior to achieve instrumental goals, such as obtaining something of value, exacting retribution for a perceived wrong, or establishing a desired social identity. (See Shaver, Segev, & Mikulincer, Chapter 4, this volume, for another functionalist perspective on aggression based on attachment theory and the concept of behavioral systems.)

In an attempt to build a broad model of aggression encompassing the common features of these theories, C. A. Anderson and Bushman (2002) honed their integrative, meta-theoretical General Aggression Model. As with I³ Theory, the General Aggression Model focuses less on a particular variable or process than on general classes of aggression risk factors and processes (DeWall & C. A. Anderson, Chapter 1, this volume). This model consists of three main foci. The first emphasizes person and situation *inputs*, or risk factors, for aggression. Person inputs include personality traits, gender, beliefs, attitudes, values, long-term goals, and scripts; situation inputs include aggressive cues (e.g., presence of guns), provocation, frustration, pain and discomfort, drugs, and incentives (determined by a cost/benefit analysis). The second focus is the

interconnected affective, arousal, and cognitive *routes*, or mechanisms, through which the inputs influence aggressive behavior. Affective routes include mood and emotion and expressive motor tendencies; arousal routes include the strengthening of a dominant action tendency or certain misattribution processes; cognitive routes include hostile thoughts and scripts. Finally, the third focus is the *outcomes* of the underlying appraisal and decision processes. Individuals are likely to act impulsively when they lack the resources and motivation to alter their immediate appraisal of the situation. If they possess the resources and motivation, however, they may reappraise the situation and act in a more thoughtful fashion (DeWall & C. A. Anderson, Chapter 1, this volume).

I³ Theory

I³ Theory, which is a process-oriented meta-theory designed to identify the circumstances under which a nonaggressive interaction can become an aggressive one, builds upon these previous theories, integrates them with recent research on self-regulation, and specifies the ways in which aggression risk factors interact to predict aggressive behavior.

The theory begins by posing three questions. First, does at least one individual in the interaction experience strong instigating triggers toward aggression? Second, does that individual experience strong impelling forces toward aggression? Third, does that individual experience weak forces to inhibit or override the aggressive impulses? Each affirmative answer increases the likelihood of aggressive behavior, via both a main effect and interactive effects with variables relevant to one or both of the other questions. Whereas the strength of impelling forces is determined by the collective power of the variables that cause the individual to experience an urge to aggress in response to a given instigating trigger, the strength of inhibiting forces is

determined by the collective power of the variables that cause the individual to override this aggressive urge.

In addition to the three initial questions posed above, I³ Theory poses a fourth: How do effects of variables in one category (instigating triggers, impelling forces, or inhibiting forces) interact with effects of one or more variables from the other categories to predict aggressive behavior? As presented in Table 1, answering these four questions enables scholars to identify seven key I³ Theory effects. Figure 1 (which builds upon work by Fals-Stewart, Leonard, & Birchler, 2005) illustrates how these seven effects can work together to increase or decrease the likelihood of aggressive behavior.

I³ Theory diverges from the aggression theories mentioned earlier in its central emphasis on inhibitory processes. The theory recognizes the importance of instigating triggers and impelling forces, but it argues that such factors cause individuals to enact aggressive behavior only when their collective power is stronger than the collective power of inhibitory processes. Although other theories address the importance of inhibitory processes in aggression (e.g., DeWall & C. A. Anderson, Chapter 1, this volume, as well as Denson, Chapter 6; Dodge, Chapter 10; and Mayseless & Scharf, Chapter 14), the emphasis on such processes gains new prominence with I³ Theory.

Instigating Triggers

The first stage of I³ Theory concerns the presence of one or more *instigating triggers*, which are discrete situational events or circumstances that induce rudimentary action tendencies toward physical aggression. As illustrated at the lower-left of Figure 1, impelling and inhibiting forces are irrelevant when instigating triggers are absent. Even the world's angriest, least controlled person is not aggressive all the time; some situational variable (even if it only serves

to activate a long-standing goal or grievance) is required before the person becomes aggressive. Aversive events can trigger (via automatic associative networks or cognitive appraisal processes) hostile cognitive, affective, physiological, and even motor tendencies that prime the individual to aggress (Berkowitz, 1993; DeWall & C. A. Anderson, Chapter 1, and Dodge, Chapter 10, this volume); I³ Theory suggests that certain events can also trigger aggressive tendencies driven by instrumental goals (e.g., having an individual offer you money to beat up his enemy).

Instigating triggers fall into one of two categories: dyadic and third-party. *Dyadic triggers* refer to events or circumstances that the potentially aggressive individual perceives as having originated in the target. Examples include direct provocation (Bettencourt & Miller, 1996), goal obstruction (Dollard et al., 1939), and social rejection (Leary, Twenge, & Quinlivan, 2006). *Third-party triggers* refer to events or circumstances that the potentially aggressive individual perceives as having originated in somebody other than the target. The same kind of trigger that leads to a rudimentary action tendency to aggress toward the provocateur can also do so toward a third party. For example, an individual who feels provoked or rejected may experience an instigation to aggress not only against the source of the provocation or rejection, but also (or alternatively) toward another target whom the potential perpetrator believes would make a more acceptable or desirable target (e.g., somebody who is less likely to fight back).

Impelling Forces

The second stage of I³ Theory concerns risk factors that determine the strength of the aggressive impulse experienced by the individual, either through main effects or through interactions with instigating triggers. In some situations, individuals may effortlessly shrug off (or perhaps not even notice; see Dodge, Chapter 10, this volume) an instigating trigger, experiencing virtually no impulse toward aggression. In others, individuals may react strongly to

a trigger, experiencing a powerful impulse toward aggression. *Impelling forces* refer to factors that increase the likelihood that individuals will experience an aggressive impulse in response to an instigating trigger. Individuals tend to experience more powerful aggressive impulses when impelling forces are strong than when they are weak (see dashed vs. dotted lines in Figure 1), especially to the degree that the instigating trigger is severe.

Impelling forces fall into one of four categories: evolutionary and cultural, personal, dyadic, and situational. *Evolutionary and cultural impellers* refer to features of the potentially aggressive individual's biological or cultural heritage, including evolutionary adaptations and social norms (see Cohen, Chapter 7, and Sell, Chapter 3, this volume). Examples include adaptations resulting from evolutionary pressures that provided ancestral men and women with a survival advantage for experiencing violent impulses in certain situations (Lorenz, 1966) and social norms delineating the extent to which certain instigating triggers provoke strong aggressive impulses (Nisbett & Cohen, 1996). *Personal impellers* refer to relatively stable characteristics of a given individual that differ from those of many other individuals, including personality characteristics, attitudes, beliefs, interpersonal interaction styles, or biological factors. Examples include dispositional hostility (Norlander & Eckhardt, 2005; Rhee & Waldman, Chapter 9, and Tackett & Krueger, Chapter 5, this volume), narcissism (Thomaes & Bushman, Chapter 11, this volume; Twenge & Campbell, 2003), and testosterone (Dabbs, Frady, Carr, & Besch, 1987; Van Goozen, Frijda, & Van de Poll, 1994). *Dyadic impellers* refer to characteristics of the relationship between the potential aggressor and the potential target. Examples include dissatisfaction with the amount of power one has in a relationship (Ronfeldt, Kimerling, & Arias, 1998), target-specific jealousy (Dutton, van Ginkel, & Landolt, 1996; Holtzworth-Munroe, Stuart, & Hutchinson, 1997), and feelings of vulnerability or insecurity in

the relationship (Carney & Buttell, 2005; Holtzworth-Munroe et al., 1997; Mikulincer & Shaver, Chapter 13, this volume). Finally, *situational impellers* refer to momentarily activated cognitive, affective, or physiological experiences. Examples include uncomfortable temperatures (C. A. Anderson, K. B. Anderson, Dorr, DeNeve, & Flanagan, 2000), physical pain (Berkowitz, 1998), and exposure to violent media (C. A. Anderson & Bushman, 2001; C. A. Anderson, Carnagey, & Eubanks, 2003).

Inhibiting Forces

The third stage of I³ Theory concerns risk factors that determine whether individuals will override the aggressive impulses that emerge from the instigating triggers, impelling factors, and their interaction. In some situations, individuals succumb to these impulses, engaging in aggressive behavior. In others, individuals override them in favor of nonviolent behavior.

Inhibiting forces refer to factors that increase the likelihood that individuals will override aggressive impulses rather than acting upon them. Inhibiting factors collectively determine the threshold above which aggressive impulses will manifest themselves in aggressive behavior. If the inhibiting forces are weak (lower horizontal line in Figure 1), then aggressive impulses need not be especially strong to result in aggressive behavior. If the inhibiting forces are strong (upper horizontal line in Figure 1), then aggressive impulses must be strong to result in aggressive behavior.

As with impelling forces, inhibiting forces fall into one of four categories: evolutionary and cultural, personal, dyadic, and situational. Examples of *evolutionary and cultural inhibitors* include adaptations resulting from evolutionary pressures that provided ancestral men and women with a survival advantage for overriding aggressive impulses in certain situations (Baumeister, 2005; Sell, Chapter 3, this volume) and social norms or institutions that decrease

the likelihood that individuals will act upon aggressive impulses (Cohen, Chapter 7, this volume; Guerra, Huesmann, & Spindler, 2002; Sampson, Raudenbush, & Earls, 1997). Examples of *personal inhibitors* include dispositional self-control (Finkel, DeWall, Slotter, Oaten, & Foshee, in press), executive functioning (Denson, Chapter 6, this volume; Giancola, 2000), and beliefs that enacting aggressive behavior will lead to poor outcomes for the self (Slaby & Guerra, 1988). Examples of *dyadic inhibitors* include partner empathy or perspective-taking (Richardson, Green, & Lago, 1998; Van Baardewijk, Stegge, Bushman, & Vermeiren, in press), relationship commitment (Gaertner & Foshee, 1999; Slotter, Finkel, & Bodenhausen, 2009), and relative physical size (Archer & Benson, 2008; Felson, 1996; Sell, Chapter 3, this volume). Finally, examples of *situational inhibitors* include low levels of alcohol use (Bushman & Cooper, 1990; Denson et al., 2008), self-regulatory resources that are not depleted (Denson, Chapter 6, this volume; DeWall et al., 2007; Finkel et al., in press), and plentiful cognitive processing time (Finkel et al., in press).

Reviewing the Aggression Literature from the Perspective of I³ Theory

One purpose of I³ Theory is to provide a coherent framework for categorizing aggression risk factors and examining the interplay among them. To illustrate how I³ Theory can accomplish these goals, we review key findings in the aggression literature through its lens, with a particular emphasis on interaction effects. This review is illustrative rather than exhaustive.

I³ Theory encompasses seven key effects: three main effects (instigating triggers, impelling forces, and inhibiting forces), three two-way interaction effects (Instigating Triggers × Impelling Forces, Instigating Triggers × Inhibiting Forces, and Impelling Forces × Inhibiting Forces), and one three-way interaction effect (Instigating Triggers × Impelling Forces ×

Inhibiting Forces). Table 1 lists these seven effects and provides an example of each one. We discuss examples of these effects in turn.

1. *An illustrative main effect of instigating triggers: Social rejection.* As shown in the first row of Table 1, an illustrative instigating trigger is *social rejection*, which refers to a class of interpersonal processes in which individuals feel rejected, excluded, or ostracized. Diverse lines of evidence demonstrate that individuals who experience social rejection are more aggressive than individuals who do not (Leary et al., 2006). In one study, participants who had been unambiguously rejected by a group of fellow participants administered substantially louder, more painful sound blasts to an unknown stranger than did participants who had been unambiguously accepted (Twenge, Baumeister, Tice, & Stucke, 2001, Study 5). In this study, social rejection served as a displaced instigating trigger because the target of the aggression was not a member of the group who had previously rejected the participant. Another study, which provided an in-depth analysis of all well-documented school shootings in the United States between 1995 and 2001, yielded compatible conclusions, with acute or chronic rejection preceding the shootings in 87% of the cases (Leary, Kowalski, Smith, & Phillips, 2003).

2. *An illustrative main effect of impelling forces: Testosterone.* As shown in the second row of Table 1, an illustrative impelling force is the androgen *testosterone*. Although testosterone is higher in men than in women, it predicts aggression in both sexes (Archer, Biring, & Wu, 1998; Dabbs & Hargrove, 1997; see Sapolsky, 1998). In a study of male prison inmates, testosterone levels correlated positively with crime severity; indeed, nine of the 11 inmates with the lowest testosterone levels had committed nonviolent crimes, whereas 10 of the 11 inmates with the highest testosterone levels had committed violent crimes (Dabbs et al., 1987). In another study, female-to-male transsexuals became considerably more aggressive in

the first three months of androgen injections (Van Goozen, Cohen-Kettenis, Gooren, Frijda, & Van de Poll, 1995).

3. *An illustrative main effect of inhibiting forces: Self-regulatory strength.* As shown in the third row of Table 1, an illustrative inhibiting factor is *self-regulatory strength*, which refers to the psychological resource that undergirds willful acts of self-regulation. According to the Strength Model of self-regulation (Baumeister, Gailliot, DeWall, & Oaten, 2006), all such acts depend on a unitary resource that resembles a muscle, in that its strength can be (a) temporarily diminished by self-regulatory exertions (leading to a state of “ego depletion”) but (b) bolstered over time by adhering to self-regulatory bolstering “regimens.” In one study, hungry participants who had previously resisted eating a tempting food (a donut; high ego depletion condition) were more aggressive toward a same-sex provoking interaction partner (forcing this partner to eat a snack with plentiful hot sauce despite his or her distaste for spicy foods) than were participants who had previously resisted eating a less tempting food (radishes; low ego depletion condition), even though participants in the two conditions did not differ in how angry they were in response to the provocation (DeWall et al., 2007, Study 1).

Complementing this evidence that ego depletion predicts elevated aggression, a recent study demonstrated that individuals who had adhered to a two-week self-regulatory strength bolstering regimen declined significantly from before to after the regimen in their aggressive tendencies toward their romantic partner (Finkel et al., in press, Study 5). In this study, participants who deliberately regulated either their physical behavior (e.g., brushing their teeth with their nondominant hand) or their verbal behavior (e.g., making sure not to begin sentences with “I”) reported a reduced likelihood of being physically aggressive in response to various

partner provocations (e.g., “I walk in and catch my partner having sex with someone”), whereas participants in a no-intervention control condition exhibited no change from pretest to posttest.

4. *An illustrative Instigating Trigger × Impelling Forces interaction effect: Ego Threat × Narcissism.* We now turn from of I³ Theory’s three main effects to its three two-way interaction effects. As shown in the fourth row of Table 1, an illustrative Instigating Trigger × Impelling Forces interaction effect is *Ego Threat × Narcissism*. Although many scholars have suggested that low self-esteem causes aggression, others have increasingly argued that a form of high self-esteem is frequently more likely to do so. In particular, individuals whose self-views are not only favorable, but also unstable, inflated, or uncertain, are especially prone toward aggressive behavior when their favorable self-views are socially threatened (Baumeister, Smart, & Boden, 1996; Thomaes & Bushman, Chapter 11, this volume). One series of studies demonstrated that participants who had experienced an ego threat in the form of insulting feedback about an essay they had written (an instigating trigger) were more aggressive toward the same-sex provocateur (subjecting him or her to painfully loud noise blasts) than were participants who had not experienced an ego threat (Bushman & Baumeister, 1998). The key finding, however, was that this main effect of ego threat was substantially larger for participants who were high in narcissism (an impelling factor) than for participants who were low.

Additional research has examined how self-views moderate the link between other instigating triggers (aside from insults) and aggression. One study demonstrated that the link between social rejection and displaced aggression (painful noise blasts) was substantially stronger for individuals who were high in narcissism than for those who were low (Twenge & Campbell, 2003, Study 4), and another demonstrated that the link between social rejection and aggression (aversive hot sauce) was substantially stronger for individuals who were high in

rejection sensitivity (who anxiously expect, readily perceive, and overreact to rejection) than for those who were low (Ayduk, Gyurnak, & Luersson, 2008).

5. *An illustrative Instigating Trigger × Inhibiting Forces interaction effect: Provocation × Self-Regulatory Strength.* As shown in the fifth row of Table 1, an illustrative Instigating Trigger × Inhibiting Forces interaction effect is *Provocation × Self-Regulatory Strength*. As discussed previously, both provocation and self-regulatory strength predict aggression via main effects. Recent studies confirm the I³ Theory prediction that incorporating their interaction effect yields a richer story (e.g., Denson, Chapter 6, this volume). Two recent experiments, one in which the aggression was directed at strangers (aversive sound blasts; DeWall et al., 2007, Study 2) and one in which the aggression was directed at one's romantic partner (forcing him or her to maintain body poses for painfully long durations; Finkel et al., in press, Study 4), demonstrated that participants were especially aggressive when they experienced *both* provocation (in the form of insulting feedback) and ego depletion.

Another example of an Instigating Trigger × Inhibiting Forces interaction effect is *Provocation Severity × Frontal Lobe Functioning* (Lau, Pihl, & Peterson, 1995). In this study, participants were preselected if they were in the upper or lower quartile on frontal-lobe-based cognitive functioning, which underlies the ability to control one's impulses (Hecaen & Albert, 1978). Consistent with I³ Theory, participants were more aggressive (administering painful electric shocks) to the degree that the opponent had previously provoked them severely rather than mildly (i.e., had administered painful electric shocks to them), but this provocation main effect was substantially stronger among individuals with weak rather than strong frontal lobe functioning.

Yet another example of an Instigating Trigger \times Inhibiting Forces interaction effect is *Provocation Salience \times Alcohol Consumption* (Densen et al., 2008). In this study, participants who were assigned to consume four alcoholic or placebo beverages in a 20-minute period were provoked in either a salient or a subtle way and then had the opportunity to aggress against their provocateur by determining for how long he or she would have to keep his or her hand immersed in iced water. Participants were more aggressive in the salient provocation condition than in the subtle provocation condition, and this effect was significantly stronger in the alcohol than in the placebo condition.

A final example of an Instigating Trigger \times Inhibiting Forces interaction effect is *Provocation Level \times Relationship Commitment* within the context of an ongoing romantic relationship (Slotter et al., 2009). In these studies, participants were more aggressive toward their romantic partner after they had been provoked by him or her, and this provocation main effect was especially strong among participants who were low in relationship commitment (and even dropped to nonsignificance among participants who were high). Extending work suggesting that commitment promotes prorelationship behaviors in other conflictual relationship domains (such as forgiveness; Finkel, Rusbult, Kumashiro, & Hannon, 2002), it appears that individuals who are highly committed to their romantic relationships are able to override aggressive impulses when their partner provokes them.

6. *An illustrative Impelling Forces \times Inhibiting Forces interaction effect: Physical Proclivity \times Negative Outcome Expectancies.* As shown in the sixth row of Table 1, an illustrative Impelling Forces \times Inhibiting Forces interaction effect is *Physical Proclivity \times Negative Outcome Expectancies*. Individuals vary in the degree to which they prefer physical versus cognitive tasks, and a relative preference for the former predicts increased tendencies

toward aggression and criminal behavior (Gottfredson & Hirschi, 1990). Individuals also vary in the degree to which they believe that engaging in aggressive behavior will cause them to experience negative outcomes, such as physical harm or social derision, and stronger beliefs in this causal link predict decreased tendencies toward aggression (Slaby & Guerra, 1988). One recent study examined whether physical proclivity (an impelling factor) interacts with negative outcome expectancies (an inhibiting factor) to predict aggression toward a romantic partner (Finkel & Foshee, 2009). Results revealed a strong positive association of physical proclivity with self-reported aggressive behavior over the previous year, but only when negative outcome expectancies were low. It seems that the tendencies to prefer physical to cognitive tasks predicts greater aggression among individuals whose aggression is not restrained by beliefs that being aggressive will bring about negative effects for them, but not among individuals whose aggression is so restrained.

7. Instigating Triggers × Impelling Forces × Inhibiting Forces interaction effects: A null set. Although the three main effects and the three two-way interaction effects reviewed above are key components of I³ Theory, the Instigating Triggers × Impelling Forces × Inhibiting Forces three-way interaction effect arguably represents the most conceptually important component of the theory. Indeed, the theory suggests that all three of the preceding two-way interaction effects are moderated by third variables from whichever category is not represented in that two-way interaction.

Testing for such a three-way interaction would not be difficult. For example, it would be easy to examine (a) whether the Ego Threat × Narcissism (Instigating Trigger × Impelling Factor) interaction effect is moderated by an inhibiting factor (e.g., self-regulatory strength, alcohol consumption, strong relationship commitment), (b) whether the Provocation × Self-

Regulation Strength (Instigating Trigger \times Inhibiting Factor) interaction effect is moderated by an impelling factor (e.g., testosterone, physical proclivity, dispositional anger), and (c) whether the Physical Proclivity \times Negative Outcome Expectancy (Impelling Factor \times Inhibiting Factor) interaction effect is moderated by an instigating trigger (e.g., social rejection, ego threat, provocation). Testing for such three-way interaction effects is an important direction for future research.

Discussion

I³ Theory seeks (a) to impose enhanced theoretical coherence on the vast array of aggression risk factors by identifying *how* each of them increases the likelihood of aggression (via instigating triggers, impelling forces, and/or inhibiting forces) and (b) to examine the manner in which risk factors from one category interface with those from one or both of the other categories to predict aggressive behavior. In the preceding discussion (also see Table 1 and Figure 1), we reviewed specific examples of how certain risk factors fit into I³ Theory and how they interface with variables from the other I³ Theory categories.

One important direction for future research (aside from providing the first tests of I³ Theory's Instigating Trigger \times Impelling Factor \times Inhibiting Factor three-way interaction effect) will be to develop and hone empirical procedures for classifying a given risk factor into an I³ Theory category (instigating trigger, impelling forces, and/or inhibiting forces)—or perhaps into more than one category if a given variable both increases aggressive impulses and decreases restraint. In this chapter, we have relied on theory to determine, for example, (a) that social rejection, ego threat, and provocation, are instigating triggers; (b) that testosterone, narcissism, and physical proclivity are impelling factors; and (c) that self-regulatory strength, lack of alcohol consumption, and negative outcome expectancies are inhibiting factors. One limitation of this

approach is that existing theory is in many cases not sufficiently developed vis-à-vis the I³ Theory parameters to allow for definitive classification. For example, we are reasonably confident that dispositional self-control predicts reduced aggression in large part by raising the inhibition threshold (see horizontal lines in Figure 1), thereby increasing the likelihood that individuals will override aggressive impulses. That said, perhaps such dispositional self-control also predicts reduced aggression in part by reducing impelling forces, thereby decreasing the strength of the aggressive impulse in the first place.

How might scholars employ empirical procedures to determine whether a given variable promotes aggression by increasing aggressive impulses or by decreasing restraint (or the proportion of the variance attributable to each mechanism)? One promising approach is to adapt recent developments in process dissociation paradigms used by social cognition researchers. Scholars have recently modeled behavior on laboratory tasks to discern the degree to which participants exhibit certain automatic tendencies (e.g., toward prejudice or discrimination) and also controlled tendencies that override these automatic tendencies (Payne, 2001; Sherman et al., 2008; see Jacoby, 1991). After developing empirical procedures for distinguishing impulses toward aggressive behavior from self-controlled processes that override those impulses, scholars will be able to examine the association of a given risk factor with both (a) individuals' tendencies to experience impulses to aggress and (b) their tendencies to override those impulses (Denson, Chapter 6, this volume). We predict that variables such as testosterone, narcissism, and physical proclivity will correlate positively with the automatic aggressive tendencies identified by these process dissociation procedures and negligibly with the controlled tendencies that override these automatic tendencies. In contrast, we predict that variables such as self-regulatory strength, sobriety, and negative outcome expectancies will correlate negligibly with the automatic

aggressive tendencies identified by these process dissociation procedures and positively with the controlled tendencies that override these automatic tendencies.

Once scholars determine (using theoretical tools, empirical tools, or both) which risk factors function by strengthening aggressive impulses and which function by weakening behavioral override processes, I³ Theory may hold promise for interventions designed to reduce aggressive behavior. For example, the theory highlights the importance of inhibitory factors (especially self-regulation) in predicting aggressive behavior, and scholars have suggested that interventions designed to help individuals override their impulses are likely to be more effective than interventions designed to prevent them from experiencing those impulses in the first place (e.g., Baumeister, 2005). To the degree that such scholars are correct, interventions designed to strengthen inhibitory forces may turn out to be more effective on average than interventions designed to weaken impelling forces. Early research inspired by I³ Theory suggests that inhibition-relevant interventions at the dispositional level (e.g., building self-regulatory strength via bolstering regimens), relational level (e.g., strengthening relationship commitment), and situational level (e.g., having participants wait 10 seconds before responding to a provocation) can reduce aggressive behavior (Finkel et al., in press; Slotter et al., 2009; see Finkel et al., 2002, for evidence that relationship commitment is amenable to experimental manipulation). And, as discussed above, distal inhibitory factors such as prevalent social norms also seem to influence aggressive behavior (Cohen, Chapter 7, this volume; Guerra, Huesmann, & Spindler, 2002; Sampson, Raudenbush, & Earls, 1997), which hints at the possibility that large-scale social interventions could potentially reduce aggression at the societal level.

Shifting back to the theoretical domain, I³ Theory is broadly consistent with all of the major theories reviewed above, agreeing that aggression is predicted by negative affect

(Berkowitz, 1993), misattributed physiological arousal (Zillman, 1983), social learning of aggressive conduct (Bandura, 1973), aggressive scripts (Huesmann et al., Chapter 8, this volume), and certain instrumental goals (Tedeschi & Felson, 1994). I³ Theory seeks to establish whether (and the degree to which) these processes promote aggression by strengthening aggressive impulses or by weakening inhibitions against aggression (e.g., misattributed physiological arousal may be more likely to predict aggressive behavior by bolstering aggressive impulses than by impairing restraint processes). To be sure, scholars in these established theoretical traditions have discussed related issues (e.g., Berkowitz, 1993); I³ Theory seeks to elevate this discussion to a primary emphasis, while simultaneously linking ideas across theories.

In conclusion, I³ Theory is an attempt to categorize aggression risk factors into instigating triggers, impelling factors, and/or inhibiting factors and to identify the interplay among variables across categories (see Figure 1). To the degree that extant theory provides good reason to believe that particular risk factors fit relatively neatly into one of the I³ Theory categories, the theory provides an immediately accessible agenda for future research, oriented less toward identifying additional risk factors than toward identifying (a) the processes by which risk factors, considered in isolation, increase aggression; and (b) the manner in which they interact to do so. In the long-run, I³ Theory can inform interventions designed to help individuals manage their aggressive impulses in a constructive manner. Indeed, scholarship may well progress to the point where interventions can be tailored to the specific inhibiting risk factors most relevant to a given person, perhaps reducing one individual's aggression by bolstering self-regulatory strength and reducing another individual's aggression by bolstering empathy. Such tailored interventions, which would require valid assessment instruments, hold particular promise for reducing the prevalence and severity of interpersonal aggression.

References

- Anderson, C. A., Anderson, K. B., Dorr, N., DeNeve, K. M., & Flanagan, M. (2000). Temperature and aggression. *Advances in Experimental Social Psychology*, *32*, 63-133.
- Anderson, C. A., & Bushman, B. J. (2001). Effects of violent video games on aggressive behavior, aggressive cognition, aggressive affect, physiological arousal, and prosocial behavior: A meta-analytic review of the scientific literature. *Psychological Science*, *12*, 353-359.
- Anderson, C. A., & Bushman, B. J. (2002). Human aggression. *Annual Review of Psychology*, *53*, 27-51.
- Anderson, C. A., Carnagey, N. L., & Eubanks, J. (2003). Exposure to violent media: The effects of songs with violent lyrics on aggressive thoughts and feelings. *Journal of Personality and Social Psychology*, *84*, 960-971.
- Archer, J., & Benson, D. (2008). Physical aggression as a function of perceived fighting ability and provocation: An experimental investigation. *Aggressive Behavior*, *34*, 9-24.
- Archer, J., Birring, S. S., & Wu, F. C. W. (1998). The Association between testosterone and aggression among young men: Empirical findings and a meta-analysis. *Aggressive Behavior*, *24*, 411-420.
- Ayduk, O., Gyurak, A., & Luerssen, A. (2008). Individual differences in the rejection-aggression link in the hot sauce paradigm: The case of rejection sensitivity. *Journal of Experimental Social Psychology*, *44*, 775-782.
- Bandura, A. (1973). *Aggression: A social learning analysis*. Englewood Cliffs: NJ: Prentice-Hall.

- Baumeister, R. F. (2005). *The cultural animal: Human nature, meaning, and social life*. New York: Oxford University Press.
- Baumeister, R. F., Gailliot, M., DeWall, C. N., & Oaten, M. (2006). Self-regulation and personality: How interventions increase regulatory success, and how depletion moderates the effects of trait on behavior. *Journal of Personality, 74*, 1773-1802.
- Baumeister, R. F., Smart, L., & Boden, J. M. (1996). Relation of threatened egotism to violence and aggression: The dark side of high self-esteem. *Psychological Review, 103*, 5-33.
- Berkowitz, L. (1993). *Aggression: Its causes, consequences, and control*. New York: McGraw-Hill.
- Berkowitz, L. (1998). Affective aggression: The role of stress, pain, and negative affect. In R. G. Geen & E. Donnerstein (Eds.), *Human aggression: Theories, research, and implications for social policy* (pp. 49-72). San Diego, CA: Academic Press.
- Bettencourt, B. A., & Miller, N. (1996). Gender differences in aggression as a function of provocation: a meta-analysis. *Psychological Bulletin, 119*, 422-447.
- Bushman, B. J. & Baumeister, R. F. (1998). Threatened egotism: Narcissism, self-esteem, and direct or displaced aggression: Does self-love of self-hate lead to violence? *Journal of Personality and Social Psychology, 75*, 2219-229.
- Bushman, B. J., & Cooper, H. M. (1990). Alcohol and human aggression: An integrative research review. *Psychological Bulletin, 107*, 341-354.
- Carney, M. M., & Buttell, F. P. (2005). Exploring the relevance of attachment theory as a dependent variable in the treatment of women mandated into treatment for domestic violence offenses. *Journal of Offender Rehabilitation, 41*, 33-61.

- Dabbs, J. M., Jr., Frady, R. L., Carr, T. S., & Besch, N. F. (1987). Saliva testosterone and criminal violence in young adult prison inmates. *Psychosomatic Medicine*, *49*, 174-182.
- Dabbs, J. M., Jr., & Hargrove, M. F. (1997). Age, testosterone, and behavior among female prison inmates. *Psychosomatic Medicine*, *59*, 477-480.
- Denson, T. F., Aviles, F. E., Pollock, V. E., Earleywince, M., Vasquez, E. A., & Miller, N. (2008). The effects of alcohol and the salience of aggressive cues on triggered displaced aggression. *Aggressive Behavior*, *32*, 25-33.
- DeWall, C. N., Baumeister, R. F., Stillman, R. G., & Gailliot, M. T. (2007). Violence restrained: Effects of self-regulation and its depletion on aggression. *Journal of Experimental Social Psychology*, *43*, 62-76.
- Dollard, J., Doob, L. W., Miller, N. E., Mowrer, O. H., & Sears, R. R. (1939). *Frustration and aggression*. New Haven, CT: Yale University Press.
- Dutton, D. G., van Ginkel, C., & Landolt, M. (1996). Jealousy, intrusiveness and intimate abusiveness. *Journal of Family Violence*, *11*, 411-423.
- Felson, R. B. (1996). Big people hit little people: Sex differences in physical power and interpersonal violence. *Criminology*, *34*, 433-452.
- Fals-Stewart, W., Leonard, K. E., & Birchler, G. R. (2005). The occurrence of male-to-female intimate partner violence on days of men's drinking: The moderating effects of antisocial personality disorder. *Journal of Consulting and Clinical Psychology*, *73*, 239-248.
- Finkel, E. J. (2007). Impelling and inhibiting forces in the perpetration of intimate partner violence. *Review of General Psychology*, *11*, 193-207.

- Finkel, E. J. (2008). Intimate partner violence perpetration: Insights from the science of self-regulation. In J. P. Forgas & J. Fitness (Eds.), *Social relationships: Cognitive, affective, and motivational processes* (pp. 271-288). New York: Psychology Press.
- Finkel, E. J., DeWall, C. N., Slotter, E., B., Oaten, M., & Foshee, V. A. (in press). Self-regulatory failure and Intimate Partner Violence perpetration. *Journal of Personality and Social Psychology*.
- Finkel, E. J. & Foshee, V. A. (2009). *Impelling and inhibiting forces in the perpetration of intimate partner violence*. Unpublished manuscript, Northwestern University.
- Finkel E. J., Rusbult, C. E., Kumashiro, M., & Hannon, P. A. (2002). Dealing with betrayal in close relationships: Does commitment promote forgiveness? *Journal of Personality and Social Psychology*, 82, 956-974.
- Gaertner, L., Foshee, V. (1999). Commitment and the perpetration of relationship violence. *Personal Relationships*, 6, 227-239.
- Giancola, P. R. (2000). Executive functioning: A conceptual framework for alcohol-related aggression. *Experimental and Clinical Psychopharmacology*, 8, 576-597.
- Gottfredson, M. R., & Hirschi, T. (1990). *A general theory of crime*. Stanford, CA: Stanford University Press.
- Guerra, N. G., Huesmann, L. R., & Spindler, A. (2002). Community violence exposure, social cognition, and aggression among urban elementary school children. *Child Development*, 74, 1561-1576.
- Hecaen, H. & Albert, M. L. (1978). *Human Neuropsychology*. New York: Wiley.

- Holtzworth-Munroe, A., Stuart, G. L., & Hutchinson, G. (1997). Violent versus nonviolent husbands: Differences in attachment patterns, dependency, and jealousy. *Journal of Family Psychology, 11*, 314-331.
- Huesmann, L. R. (1998). The role of social information processing and cognitive schema in the acquisition and maintenance of habitual aggressive behavior. In R. G. Geen & E. Donnerstein (Eds.), *Human aggression: Theories, research, and implications for social policy* (pp. 73-109). San Diego, CA: Academic Press.
- Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30*, 513-541.
- Lau, M. A., Pihl, R. O., & Peterson, J. B. (1995). Provocation, acute alcohol intoxication, cognitive performance, and aggression. *Journal of Abnormal Psychology, 104*, 150-155.
- Leary, M. R., Kowalski, R. M., Smith, L., & Phillips, S. (2003). Teasing, rejection, and violence: Case studies of the school shootings. *Aggressive Behavior, 29*, 202-214
- Leary, M. R., Twenge, J. M., & Quinlivan, E. (2006). Interpersonal rejection as a determinant of anger and aggression. *Personality and Social Psychology Review, 10*, 111-132.
- Lorenz, K. (1966). *On aggression*. San Diego, CA: Harcourt Brace.
- Nisbett, R. E., & Cohen, D. (1996). *Culture of honor: The psychology of violence in the South*. Boulder, CO: Westview Press.
- Norlander, B., Eckhardt, C. (2005). Anger, hostility, and male perpetrators of intimate partner violence: A meta-analytic review. *Clinical Psychology Review, 25*, 119-152.
- Payne, B. K. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology, 81*, 181-192.

- Richardson, D. R., Green, L. R., & Lago, T. (1998). The relationship between perspective-taking and nonaggressive responding in the face of an attack. *Journal of Personality, 66*, 235-256.
- Ronfeldt, H. M., Kimerling, R., & Arias, I. (1998). Satisfaction with relationship power and the perpetration of dating violence. *Journal of Marriage and the Family, 60*, 70-78.
- Sampson, R. J., Raudenbush, S. W., & Earls, F. (1997). Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science, 277*, 918-924.
- Sapolsky, R. M. (1998). *The trouble with testosterone: And other essays on the biology of the human predicament*. New York: Scribner.
- Schank, R. C., & Abelson, R. P. (1977). *Scripts, Plans, Goals and Understanding: An Inquiry into Human Knowledge Structures*. Hillsdale, NJ: Erlbaum.
- Sherman, J. W., Gowronski, B., Gonsalkorale, K., Hugenberg, K., Allen, T. J., & Groom, C. J. (2008). The self-regulation of automatic associations and behavioral impulses. *Psychological Review, 115*, 314-335.
- Slaby, R. G., & Guerra, N. G. (1988). Cognitive mediators of aggression in adolescent offenders: I. Assessment. *Developmental Psychology, 24*, 580-588.
- Slotter, E. B., Finkel, E. J., & Bodenhausen, G. V. (2009). *Provocation and commitment in intimate partner violence*. Unpublished manuscript, Northwestern University.
- Tedeschi, J. T., & Felson, R. B. (1994). *Violence, aggression, and coercive actions*. Washington, DC: American Psychological Association.
- Twenge, J. M., Baumeister, R. F., Tice, D. M., & Stucke, T. J. (2001). If you can't join them, beat them: Effects of social exclusion on aggressive behavior. *Journal of Personality and Social Psychology, 81*, 1058-1069.

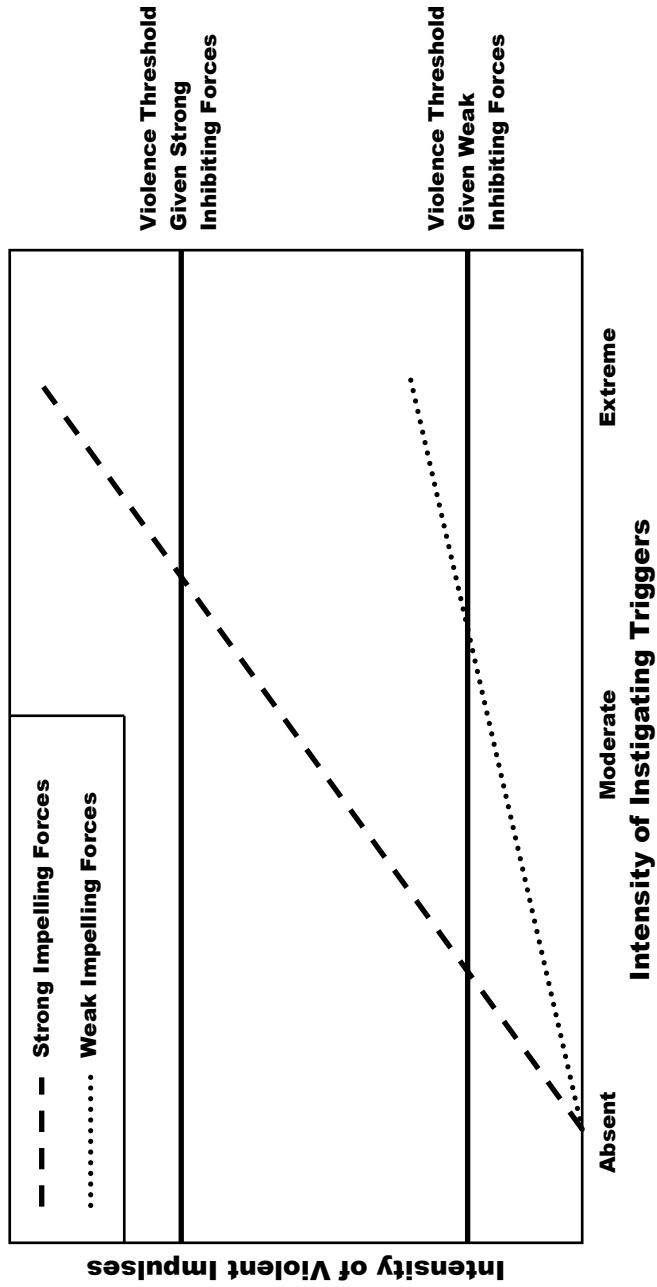
- Twenge, J. M., & Campbell, W. K. (2003). Isn't it fun to get the respect we deserve? Narcissism, social rejection, and aggression. *Personality and Social Psychology Bulletin*, 29, 261-272.
- Van Baardewijk, Y., Stegge, H., Bushman, B. J., & Vermeiren, R. (in press). "I feel your pain" (but only when I am forced to focus on it): Psychopathy, distress, and aggressive behavior. *Journal of Child Psychology and Psychiatry*.
- Van Goozen, S., Cohen-Kettenis, P. T., Gooren, F. J., Frijda, N., & Van de Poll, N. (1995). Gender differences in behaviour: Activating effects of cross-sex hormones. *Psychoneuroendocrinology*, 20, 343-363.
- Van Goozen, S., Frijda, N., & Van de Poll, N. (1994). Anger and aggression in women: Influence of sports choice and testosterone administration. *Aggressive Behavior*, 20, 213-222.
- Zillman, D. (1983). Arousal and aggression. In R. G. Geen & E. Donnerstein (Eds.), *Aggression: Theoretical and empirical reviews* (pp. 75-102, vol. 1). New York, NY: Academic Press.

Table 1

I³ Theory of Aggression: The Seven Effects

I ³ Effect (#)	I ³ Effect (Stage)	I ³ Effect (Description)	Example	Citation for Example
1	Stage 1	Instigating Trigger main effect	Social Rejection	Twenge et al., 2001
2	Stage 2	Impelling Forces main effect	Testosterone	Dabbs et al., 1987
3	Stage 3	Inhibiting Forces main effect	Self-Regulatory Strength	DeWall et al., 2007
4	Stage 1 × Stage 2	Instigating Trigger × Impelling Forces Interaction Effect	Ego Threat × Narcissism	Bushman & Baumeister, 1998
5	Stage 1 × Stage 3	Instigating Trigger × Inhibiting Forces Interaction Effect	Provocation × Self-Regulatory Strength	Finkel et al., in press
6	Stage 2 × Stage 3	Impelling Forces × Inhibiting Forces Interaction Effect	Physical Proclivity × Negative Outcome Expectancies	Finkel & Foshee, 2009
7	Stage 1 × Stage 2 × Stage 3	Instigating Trigger × Impelling Forces × Inhibiting Forces Interaction Effect	[No Known Example]	[No Known Example]

Figure 1. How the three components of \hat{P} Theory interrelate to predict aggressive behavior.



Note. For ease of illustration, Figure 1 depicts impelling forces and inhibiting forces as if they are binary—either weak or strong. In reality, the intensity of each type of force varies continuously from weak to strong.