aps ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Perspectives on Psychological Science
6(5) 415–427
© The Author(s) 2011
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1745691611416992
http://pps.sagepub.com
SAGE

Drunk, Powerful, and in the Dark: How General Processes of Disinhibition Produce Both Prosocial and Antisocial Behavior

Jacob B. Hirsh[1], Adam D. Galinsky[2], and Chen-Bo Zhong[1]
[1]University of Toronto, Canada, and [2]Northwestern University, Evanston, IL

Abstract

Social power, alcohol intoxication, and anonymity all have strong influences on human cognition and behavior. However, the social consequences of each of these conditions can be diverse, sometimes producing prosocial outcomes and other times enabling antisocial behavior. We present a general model of disinhibition to explain how these seemingly contradictory effects emerge from a single underlying mechanism: The decreased salience of competing response options prevents activation of the Behavioral Inhibition System (BIS). As a result, the most salient response in any given situation is expressed, regardless of whether it has prosocial or antisocial consequences. We review three distinct routes through which power, alcohol intoxication, and anonymity reduce the salience of competing response options, namely, through Behavioral Approach System (BAS) activation, cognitive depletion, and reduced social desirability concerns. We further discuss how these states can both reveal and shape the person. Overall, our approach allows for multiple domain-specific models to be unified within a common conceptual framework that explains how both situational and dispositional factors can influence the expression of disinhibited behavior, producing both prosocial and antisocial outcomes.

Keywords

disinhibition, BIS, power, alcohol, anonymity

What distinguishes between an amorous drunk and a hostile one looking for a fight? When does power corrupt and when does it promote social responsibility? Why does anonymity sometimes bring out the best in people and sometimes the worst? Power, alcohol intoxication, and concealed identity can have profound social consequences, but the valence and range of the outcomes is surprisingly contradictory. Experiencing high levels of social power can lead to corruption and unethical behavior in the pursuit of self-interest (Lammers, Stapel, & Galinsky, 2010), but can also result in a heightened concern for others and more generous contributions to collective resources (Chen, Lee-Chai, & Bargh, 2001; Galinsky, Gruenfeld, & Magee, 2003). Intoxicated individuals can be both more aggressive (Pihl, Peterson, & Lau, 1993) and altruistic (Steele, Critchlow, & Liu, 1985) than people who are sober. Finally, anonymity can increase selfishness and cheating (Zhong, Bohns, & Gino, 2010), but it can also promote helping behavior (Bohns, Gino, & Zhong, 2010).

Although each of these research streams is studied independently from the others, the current article describes how all of these contradictory effects can emerge from a single underlying mechanism. In particular, we draw upon Jeffrey Gray's research on the Behavioral Inhibition System to offer a general model of disinhibition that focuses on the reduced salience of competing response options to account for these diverse and seemingly contradictory effects. This model of disinhibition can also explain how power, alcohol, and anonymity both reveal the person, leading to greater correspondence between underlying dispositions and behavior, and shape the person by leading individuals to behave more consistently with strong situational cues.

Mechanisms of Disinhibition

In his seminal book on the neuropsychology of anxiety, Jeffrey Gray outlined the functional properties of a brain system involved in the slowing or cessation of ongoing behavior in response to unexpected events or cues for punishment (Gray, 1982). Using a combination of behavioral, pharmacological, and electrophysiological methods, he identified the septal-hippocampal system as the neural basis of behavioral expressions of anxiety and the primary target of anxiolytic drugs. Gray labeled this brain network the Behavioral Inhibition

Corresponding Author:
Jacob B. Hirsh, University of Toronto, Rotman School of Management, 105 St. George Street, Toronto, Ontario, M5S 3E6, Canada
E-mail: jacob.hirsh@rotman.utoronto.ca

System (BIS) and proposed that it serves to interrupt any behaviors that may lead to aversive consequences, producing a heightened state of arousal and a vigilant scanning of the environment to determine the correct course of action. Gray's research has had a tremendous influence and is featured prominently in a number of models of behavioral inhibition (Carver & White, 1994; Elliot & Thrash, 2002; Patterson & Newman, 1993; Sutton & Davidson, 1997).

Importance of goal conflict and routes to disinhibition

Although Gray's earlier research characterized the BIS as being responsive primarily to threats or cues for punishment, his later work emphasized the importance of goal-conflict (Gray & McNaughton, 2000). In particular, BIS activation is said to arise when competing motor responses are simultaneously activated—for example, when one goal calls for a certain action and a second goal simultaneously calls for a conflicting action. Thus, BIS activity occurs not only in the face of potential threats, but also when the appropriate response to a given situation is not clear (McNaughton & Gray, 2000). In Gray's revised model, the anxiety and attentional vigilance that results from BIS activity serves to interrupt ongoing behavior during states of goal conflict so that the most appropriate course of action can be identified.

Disinhibition in this model refers to a state in which the relative strength of any competing motor response is decreased, thereby allowing the most salient action to be expressed without interference from the BIS. States in which there are multiple competing responses, by contrast, activate the BIS and are characterized by inhibition of behavior and a heightened experience of anxiety (Emmons & King, 1988; Frone, 2000; Gray & McNaughton, 2000).

More recently, researchers have implicated the anterior cingulate cortex (ACC) in conflict-detection processes (Botvinick, Braver, Barch, Carter, & Cohen, 2001; Yeung, Botvinick, & Cohen, 2004). In particular, situations involving high levels of response conflict tend to be characterized by heightened levels of ACC activity, which in turn engages attentional control mechanisms in the dorsolateral prefrontal cortex to determine the best course of action (A. W. MacDonald, Cohen, Stenger, & Carter, 2000). The ACC appears to act as a cortical extension of the septo-hippocampal BIS system, helping to adapt behavioral responses as environmental circumstances change (Luu, Tucker, Derryberry, Reed, & Poulsen, 2003). Increased ACC activity has also been associated with greater anxiety (Hajcak, McDonald, & Simons, 2003), as well as higher scores on dispositional measures of BIS sensitivity (Amodio, Master, Yee, & Taylor, 2008; Boksem, Tops, Wester, Meijman, & Lorist, 2006). As predicted by Gray's model of behavioral inhibition, individuals with a reduced ability to detect conflict between competing response cues also tend to be more impulsive in their behavior—when conflict between competing response options remains undetected, the BIS is not called onto active duty (Olvet & Hajcak, 2008). More generally, disinhibited states appear to emerge when a single dominant response option is perceived, preventing conflict-related activity in the BIS from inhibiting prepotent behaviors.

Applying general processes of disinhibition to three diverse domains

Adopting this conflict-based perspective on disinhibition allows seemingly distinct research literatures to be united by a common mechanism. Specifically, we argue that the paradoxical effects of social power, alcohol, and anonymity are all related to a decreased salience of competing response options, which in turn results in less conflict-related BIS activity. Although the BIS framework can be applied to many different phenomena, we focus on these specific subject areas for two reasons. First, they appear to be unrelated to one another on the surface, thereby demonstrating the breadth of a general model of disinhibition. Second, these examples illustrate three distinct pathways to a disinhibited state: Social power results in a heightened approach motivation and goal focus, alcohol intoxication impairs attentional resources, and anonymity reduces social desirability concerns. Although the causal pathways are different in each of these cases, we argue that they all result in a disinhibited state characterized by reduced response conflict. Consequently, there is no inhibitory interference from the BIS and the most salient response in any given situation is likely to be enacted, regardless of whether it has prosocial or antisocial consequences. Figure 1 highlights the three distinct paths to disinhibition that are examined in this article, each characterized by a decreased salience of competing response options and a consequent reduction in BIS activity.

It is important to note that the most salient response can be internally triggered by an individual's chronically active motives and dispositions. This is especially true in weak situations without strong contextual norms and cues. Thus, disinhibition can reveal the person, leading behavior to be more consistent with one's underlying tendencies. However, the most salient response can also be externally triggered when there are strong cues present in the environment. When situational cues dominate, disinhibition shapes the person by leading them to act in line with prominent environmental affordances. In contexts where neither the situation nor the disposition is clearly stronger, the most salient responses are likely to be those that receive joint input from both internal and external cues.

In this article, we review the three routes to disinhibition via social power, intoxication, and anonymity, demonstrating how the salience of competing response options is reduced in each case. We further emphasize how the effects of both the individual's dispositional characteristics and the situation's most salient cues can be enhanced via these disinhibitory effects to produce both prosocial and antisocial behaviors.

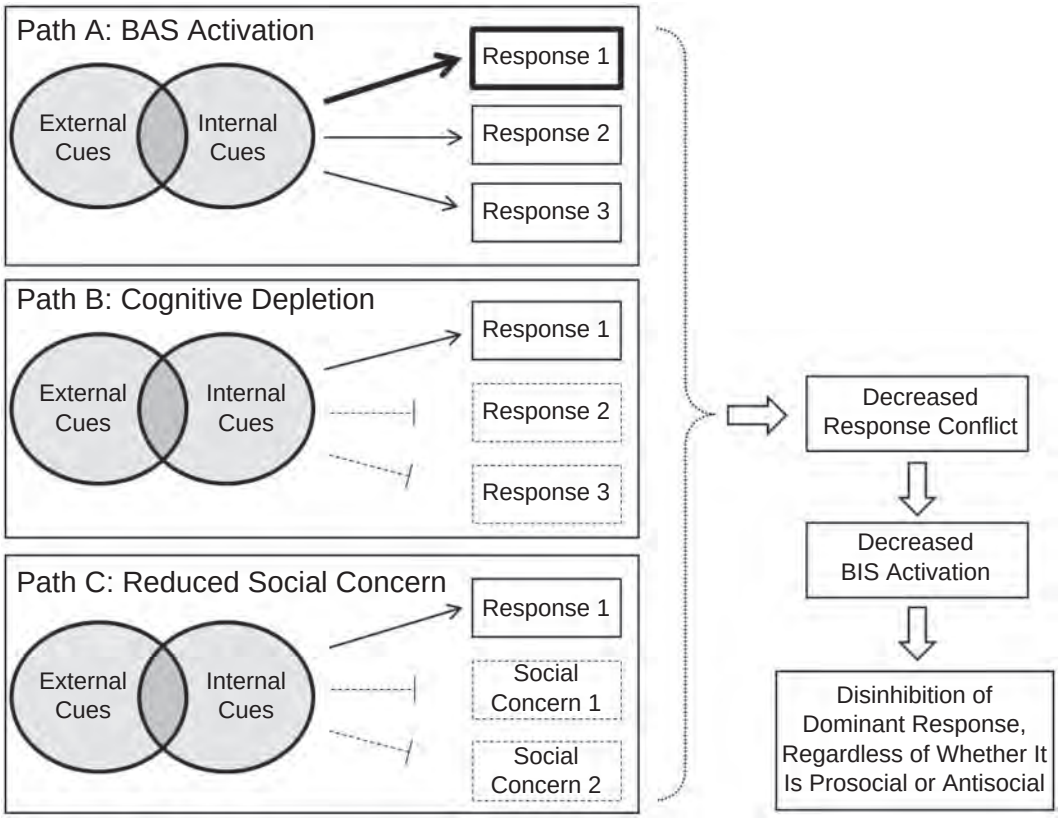


Fig. 1. Three pathways to disinhibition. In Path A, activity in the Behavioral Approach System (BAS) increases the strength of the most salient response. In Path B, depletion of cognitive resources hinders the perception of nonsalient response cues. In Path C, reduction in social concern lowers the activation of any responses derived from evaluative concerns. Each pathway reduces the simultaneous activation of competing responses, thereby decreasing activity in the Behavioral Inhibition System (BIS) and producing only a single dominant response (i.e., disinhibition). The most salient response will in turn depend upon the strength of internal cues (e.g., dispositions) and external cues (e.g., situational norms); the meaning and implications of these internal and external cues may sometimes be congruent with one another (as represented by their partially overlapping circles). Disinhibition can result in antisocial or prosocial consequences, depending on the valence of the most salient response option.

Disinhibition Through BAS Activation: The Case of Social Power

The first mechanism by which response conflict can be reduced is through heightened activity in the Behavioral Approach System (BAS), a dopaminergically-mediated brain circuit associated with the approach and pursuit of potential rewards. The BAS is thought to support appetitive movement toward a desired goal, and it is often associated with positive affect and forward locomotion (Gray, 1990; Watson, Wiese, Vaidya, & Tellegen, 1999). An important cognitive consequence of BAS activation is that attention tends to narrow towards the desired goal (Gable & Harmon-Jones, 2008, 2010). Behaviorally, this means that any actions associated with the currently active goal are strengthened, whereas any competing behavioral responses receive less activation (Harmon-Jones & Harmon-Jones, 2002; Shah, Friedman, & Kruglanski, 2002).

Because of this selective enhancement of goal-related behaviors, strong engagement of the BAS reduces the salience of competing responses and thus results in less conflict-related BIS activity during response selection, allowing goal-directed behavior to be expressed in an uninhibited manner. More generally, the BAS and BIS systems tend to have an antagonistic relationship, such that increasing activation in one system decreases activation in the other (Corr, 2002, 2004). For this reason, people often engage in the zealous pursuit of goal-directed behavior to reduce BIS-related anxiety through heightened BAS activity (McGregor, Nash, Mann, & Phills, 2010).

This BAS-related silencing of the BIS suggests that any situation that heightens BAS activity is likely to lead to a disinhibited state in which a single dominant response influences behavioral output. Some of the factors that increase BAS activity include sexual arousal (Janssen, Vorst, Finn, & Bancroft, 2002), approach-oriented affect (Carver, Sutton, & Scheier, 2000), and high levels of extraversion (Depue & Collins, 1999). In this section, we focus on an additional pathway to increased BAS-related disinhibition: social power.

Social power refers to asymmetric control over valued resources in social relations (Emerson, 1962; Magee &

Galinsky, 2008). Variations in the experience of power (or lack thereof) can have profound psychological effects, resulting in a number of cognitive and behavioral differences between the powerful and the powerless (Guinote & Vescio, 2010). By definition, the powerful live in a world of relative abundance, such that they have increased access to resources and potential rewards. As a result, social power tends to be associated with heightened activity in the BAS, as evidenced by increased left-frontal brain activity after an individual is primed with power (Boksem, Smolders, & De Cremer, 2009). This greater BAS activation among the powerful then promotes approach toward potential rewards (Inesi, 2010; Keltner, Gruenfeld, & Anderson, 2003; Smith & Bargh, 2008).

One of the defining psychological consequences of having power—an increased goal focus—is also consistent with heightened BAS activity (Galinsky et al., 2003; Guinote, 2007b). Indeed, the wide range of effects of power can all be united and parsimoniously explained through an increased focus on goals and facilitation of goal-directed behavior (Galinsky, Rus, & Lammers, 2010). The source of this goal-directed behavior can be internally driven by dispositional motives or externally driven by situational cues.

Power reveals the person: The influence of dispositional factors

As goal-focused BAS activity increases with the experience of power, there is less inhibition from competing responses to restrain the expression of personal motives. Indeed, numerous studies have found that power reveals the person by increasing the correspondence between traits and behavior, with the personalities of high-power individuals being better predictors of their behavior than the personalities of low-power individuals. For example, individuals with a communal orientation are more likely to behave in socially responsible ways when primed with power, whereas those with an exchange orientation behave in self-serving manners following the same manipulation (Chen et al., 2001). In one Chen et al. study, communally oriented participants were willing to work longer on an experimental task when primed with power, thereby decreasing the amount of labor required from their subordinates. Conversely, the power prime led exchange-oriented participants to perform less work during the same task. Similar disposition effects are reflected in the observation that priming men with power leads them to view their female work partners in sexual terms, but only for those men with a predisposition toward sexual harassment (Bargh, Raymond, Pryor, & Strack, 1995). Powerful individuals are also more likely to act in line with their preexisting value orientations when negotiating with others (Galinsky, Magee, Gruenfeld, Whitson, & Liljenquist, 2008). For high-power negotiators, social value orientation, a personality trait that describes preferences for allocations between the self and others, significantly predicted how much they trusted their opponent when entering a negotiation. This trait did not, however, predict the trust levels of nonpowerful participants. Thus, power led to divergent negotiation styles by amplifying participants' prior dispositional orientations. Related to this finding, increased power can also lead to greater interpersonal sensitivity, but only among those who identify with an empathic leadership style rather than an egoistic leadership style (Côté et al., 2011; Mast, Jonas, & Hall, 2009). Similarly, the effects of subliminally priming power appear to have diverging effects depending on one's internalized cultural values: Individuals from individualistic Western cultures who were primed with power more readily accessed words related to the concept of entitlement, whereas East Asians primed with power more readily accessed words related to social responsibility (Zhong, Galinsky, Magee, & Maddux, 2011).

Overall, social power appears to bias individuals to behave in line with their preexisting attitudes and dispositions, regardless of the social consequences (Anderson & Berdahl, 2002; Keltner et al., 2003). Power increases goal-focused BAS activity, which reduces inhibition from competing responses that would otherwise limit the expression of one's personality and personal motives.

Power shapes the person: The influence of situational factors

In addition to the influence of dispositional variables, the effects of power can also be altered by the presence of salient situational cues. Powerful individuals appear to be more responsive to the dominant affordances in any given situation, such that contextually activated goals are more likely to be pursued when in a powerful state (Guinote, 2008). The tendency for the powerful to actively engage in goal-directed behavior can thus lead to diverse outcomes depending upon the nature of the social situation. For instance, public-goods and commons dilemmas are monetary allocation games that are structurally similar to one another but involve very different actions. Action in the commons dilemma involves taking resources away from the group, whereas action in the public-goods dilemma involves giving resources to the group. Greater social power results in a greater number of actions being taken in each of these dilemmas, regardless of the social consequences of those actions (Galinsky et al., 2003). Similarly, whereas competitive contexts tend to result in more self-interested uses of power, cooperative contexts result in more generous and supportive uses of power (Handgraaf, Van Dijk, Vermunt, Wilke, & De Dreu, 2008; Tjosvold, 1985). When the powerful have tempting rewards dancing in front of them, they are more likely to cheat to capture these tantalizing possibilities (Lammers et al., 2010). However, when the situation involves strong cues for altruism, such as cries for help, the powerful are more likely to intervene in an emergency (Whitson, Liljenquist, Galinsky, Magee, & Gruenfeld, 2011). The situational context can thus establish the salient goal that powerful individuals pursue, thereby influencing whether their disinhibited behavior will result in prosocial or antisocial outcomes.

Power summary

The panoply of power findings we have reviewed is consistent with our broad conceptualization that disinhibition occurs when the salience of alternative response options is reduced. In any given situation, an individual must choose the appropriate action from an array of competing behavioral affordances. Response conflict and behavioral inhibition occur when two or more affordances are equally salient, resulting in ambiguous perception of the appropriate action (Gray & McNaughton, 2000).

We have proposed that powerful individuals experience less response conflict due to heightened BAS-related activity, which narrows goal-focused attention. For example, in a public goods dilemma, the most salient action of contributing resources to the group is often inhibited by the competing tendency to distrust the intentions of others and refrain from giving. Power helps to reduce the salience of this competing response, so that the powerful are more likely to follow the most salient response option and contribute to the group without hesitation (Galinsky et al., 2003). Regardless of whether the dominant response emerges from a person's disposition or the situation, power disinhibits both prosocial and antisocial behavior by reducing the salience of competing response options. Through the general processes of disinhibition, power both reveals and shapes the person.

Disinhibition Through Cognitive Depletion: The Case of Alcohol Intoxication

A second pathway by which response conflict and associated BIS activity can be decreased is through the depletion of cognitive resources. Our finite cognitive capacities limit the amount of information that we are able to attend to in any given moment (Baddeley, 1986). Although we are continuously presented with a large number of perceptual cues and behavioral affordances, we can only focus on a relatively small number of these at any given time. The availability of attentional resources thus affects the ability to perceive and encode competing response options (Awh, Vogel, & Oh, 2006; Engle, Kane, & Tuholski, 1999). During moments in which more cognitive resources are available, a broader array of response cues and potential actions can be perceived. When cognitive resources are more limited, however, only the most salient responses are likely to be attended to and pursued.

Consequently, the reduction of cognitive resources is likely to lead to a state of disinhibition, as the salience of competing response options is lowered and conflict-related BIS activity is correspondingly decreased. Some examples of this pathway to disinhibition include dispositionally lower working-memory capacities (Engle, 2002), heightened cognitive load or distraction (Ward & Mann, 2000), aging (von Hippel, 2007), and exhaustion or depletion effects (Vohs & Heatherton, 2000). In this section, we focus on the specific domain of alcohol intoxication.

The psychological effects of alcohol intoxication are often understood in terms of *alcohol myopia*, a state in which attentional capacities are reduced and shallower cognitive processing occurs (Steele & Josephs, 1990; Steele & Southwick, 1985). As a result, weak cues for guiding behavior are no longer extensively processed. It is for this reason that intoxicated individuals tend to perform worse in tasks in which a prepotent response must be inhibited in favor of an alternative response such as in the go–stop task or the Stroop task (Curtin & Fairchild, 2003; Easdon & Vogel-Sprott, 2000; Fillmore & Vogel-Sprott, 1999). Because weaker, less focal cues receive less attention, there is less activation of conflicting responses and therefore less BIS-related inhibition of any dominant behavioral tendencies. As a result, the most salient cues in any given moment are more likely to guide behavior when intoxicated. In keeping with Gray's model, alcohol also appears to directly decrease the responsivity of the septo-hippocampal system, thereby decreasing the experience of anxiety and the expression of behavioral inhibition (Gray & McNaughton, 2000; Peterson, Pihl, Seguin, Finn, & Stewart, 1993).

As in our general model of disinhibition, inebriation can both reveal and shape the person. The most salient cues for guiding drunken behavior emerge from both internal and external sources, regardless of whether they produce prosocial or antisocial outcomes.

Inebriation reveals the person: The influence of dispositional factors

With regard to dispositional influences, preexisting attitudes and personality characteristics tend to moderate the behavioral consequences of alcohol intoxication. For example, although aggression is a common result of alcohol consumption (Pihl et al., 1993), it occurs primarily among those individuals who have higher levels of dispositional aggressive tendencies to begin with (Giancola, 2002; Moeller, Dougherty, Lane, Steinberg, & Cherek, 1998). For instance, Giancola (2002) administered alcohol or placebo beverages to participants prior to the completion of what appeared to be a competitive reaction time task (although in actuality, the opponent was computer controlled). Participants were told that if they were the fastest to respond on a given trial, they would have the opportunity to shock their opponent. Conversely, they received a shock whenever their purported opponent was faster. Aggression was measured as the intensity of shock that the participant chose to deliver to their opponent following a successful trial. Inebriation increased aggression in those who showed a high level of dispositional aggressiveness but did not do so in those who showed low levels. Among dispositionally aggressive individuals, the salient response is to shock one's opponent, whereas the conflicting response reflects concerns of nonviolence as well as the potential for retaliation. Because alcohol reduces the attention that is allocated to such conflicting response cues (Steele & Josephs, 1990), participants' dispositionally heightened aggressive impulses were able to be

expressed more freely when drunk. Conversely, individuals with high levels of dispositional empathy are less likely to behave aggressively when intoxicated, which is consistent with their preexisting motivational profile reflecting a heightened sympathy for others (Giancola, 2003). Risky sexual decision making, another common effect of alcohol consumption, is likewise moderated by individuals' initial beliefs about the dangers of such actions (Davis, Hendershot, & George, 2007). Thus, similar to social power, alcohol intoxication can also reveal the person by increasing the correspondence between underlying dispositions and behavior.

Inebriation shapes the person: The influence of situational factors

In keeping with our general model of disinhibition, the effects of alcohol can also be influenced by salient situational cues. For instance, the provision of an explicit norm of nonaggression has the effect of reducing aggressive behavior while intoxicated (Jeavons & Taylor, 1985). In the Jeavons and Taylor study, inebriated individuals administered shocks that were 75% less intense when there was a salient norm of nonaggression than they did when there was no such norm. Drunk individuals are also more helpful than are their sober counterparts when the situation involves strong cues for helpful behavior, such as a direct request for assistance from another person needing help (Steele et al., 1985). In such cases, there appears to be less attention paid to competing cues for not helping (e.g., the potential costs of getting involved), resulting in the disinhibition of more generous behavior. Conversely, intoxicated aggression is greater in situations in which an anger response is provoked (Gustafson, 1993). By the same token, the presence of dominant situational cues can either promote or reduce risky sexual decision making amongst intoxicated individuals (Cooper, 2006; T. K. MacDonald, Fong, Zanna, & Martineau, 2000). T. K. MacDonald et al. found that when participants were saliently primed with ideas about sexual intercourse, intoxicated individuals produced riskier sexual decisions than the sober group, whereas no differences were observed following the presentation of more subtle cues (cf. Ariely & Loewenstein, 2006). However, when a salient message about the dangers of unprotected sex was present, drunken participants actually become more conservative (i.e., more likely to practice safe sex) than did the sober ones (T. K. MacDonald et al., 2000).

Alcohol summary

Overall, alcohol-induced depletion of cognitive resources allows for the processing of only the most salient response options, with competing cues and affordances receiving less attention. Sometimes the most salient response option is produced internally from dispositional characteristics, and sometimes it is produced externally from strong situational cues. In either case, a reduction in conflict-related BIS activity results in a variety of disinhibited behaviors, which can be either prosocial or antisocial depending on the specific situational and dispositional context.

Disinhibition Through Reduced Social Desirability Concerns: The Case of Anonymity

A third pathway by which the salience of competing responses can be reduced is through the reduction of social desirability concerns. Concern about social desirability is a powerful force in human society, as there can be serious adaptive consequences to being rejected by one's social group (Baumeister & Leary, 1995). During situations with relatively high levels of personal accountability, any behavioral response must be regulated in line with impression management goals in order to maintain desired levels of social acceptance (Leary & Kowalski, 1990). For instance, even though an employee may feel anger toward a coworker, the initial desire to express that anger may be inhibited by the competing goal of maintaining a positive relationship. When concerns about social evaluation are decreased, however, basic motivational impulses can be expressed more readily without being inhibited by the simultaneous and conflicting activation of socially desirable responses and the need to present a positive self-image to others (Joinson, 1999).

In this section, we focus on situations in which social desirability concerns are reduced through the concealment of personal identity, such as when in the dark, wearing a mask, or interacting in an anonymous chat room. Related to the notion of deindividuation (Diener, 1979), these contexts involve a reduction in self-awareness and the concern for projecting a positive self-image to others. Such anonymity can also serve to reduce the experience of anxiety (and therefore the BIS activity which supports it), which is consistent with our general model of disinhibition (Shepherd & Edelmann, 2005). As in other disinhibited states, a sense of anonymity can result in both higher levels of prosocial activity such as honesty and self-disclosure, as well as antisocial activities such as aggression and verbal abuse (Joinson, 2007). Similarly, anonymity, like power and alcohol, can both reveal and shape the person.

Anonymity reveals the person: The influence of dispositional factors

Whether the disinhibited behavior stemming from anonymity is "toxic" or "benign" appears in part to depend upon the preexisting disposition of the individual (Suler, 2004). Large crowds can produce a sense of anonymity and loss of personal identity in everyone (Reicher, 2001), but it is the dispositionally aggressive individuals that are most likely to escalate levels of crowd violence when the opportunity presents itself (Russell, 2004). For instance, Russell (1995) asked spectators at an ice hockey game how likely they would be to contribute to a sports riot and found a positive relationship with

aggressive personality characteristics. Conversely, less aggressive individuals in the same environment reported that they would be more likely to intervene to stop any fights that broke out (Russell & Mustonen, 1998). Consistent with research on gender differences that have found that males in general are more aggressive than females (Eagly, 1987; Eagly & Steffen, 1986), male and female groups also respond differently to anonymity and deindividuation, with only all-male groups engaging in aggressive antisocial behaviors when deindividuated (Cannavale, Scarr, & Pepitone, 1970). Similarly, people are less concerned about social evaluation when interacting over the Internet, allowing them to express themselves more directly and with less inhibition (Bargh, McKenna, & Fitzsimons, 2002). With fewer external constraints on action, anonymous situations allow internally generated motives to reveal themselves more clearly.

Anonymity shapes the person: The influence of situational factors

Strong situational cues also have a potent influence in shaping behavior during states of concealed identity. Although anonymity was once thought to exclusively produce increases in antisocial or antinormative actions (Festinger, Pepitone, & Newcomb, 1952; Zimbardo, 1969), a meta-analysis of deindividuation studies reports that anonymous individuals display an increased conformity to situation-specific norms and that they rely on salient contextual cues for guiding behavior (Postmes & Spears, 1998). In other words, the specific expression of disinhibited anonymity can be positive or negative, depending on the dominant responses afforded by strong situations and their salient norms. Although these responses might normally be inhibited by conflicting impression management goals, anonymity allows the situation to have a direct and uninhibited influence on behavioral outcomes by reducing the activation of competing response options.

Thus, anonymity can result in both prosocial and antisocial outcomes depending upon the valence of the most salient situational cue (Gergen, Gergen, & Barton, 1973; Johnson & Downing, 1979). For instance, Johnson and Downing (1979) asked participants to wear one of two costumes—either one that looked like a nurse's outfit or one that resembled a Ku Klux Klan uniform—while taking part in a study purportedly examining the effects of arousal on learning. Participants were told that they could alter the intensity of an electric shock to be administered to a student in another room whenever he made errors during the task. When anonymity was increased by making the participants unidentifiable to the purported student in the other room and the experimenter, participants in the Ku Klux Klan outfit tended to increase the intensity of the shock, whereas those in the nurse's uniform tended to decrease the shock. In this case, anonymity decreased concern about how one's actions would be evaluated (by the experimenter and other participants), allowing the situational cues to have a stronger effect on behavior.

Dim lighting also has the effect of heightening perceived anonymity, but can produce either aggression and self-interested behavior (Prentice-Dunn & Rogers, 1980; Zhong et al., 2010) or greater prosocial behavior (Bohns et al., 2010), depending on the context. For example, dimmed lighting or donning a pair of sunglasses allows individuals to feel anonymous psychologically, and this illusory anonymity increases cheating. In one Zhong et al. study, participants completed a math task in which they graded their own answers and paid themselves based on performance. Participants in a dimly lit room reported solving a greater number of problems than they actually did and thus paid themselves significantly more. In this case, participants found themselves in a dilemma in which cheating is a salient option that can produce greater rewards, but may engender negative social consequences and sanctions. In such a situation, the competing social desirability concerns normally activate the BIS system, which in turn inhibits deceptive behavior. When the sense of anonymity increases under dim lighting, however, the concern about being caught is no longer salient enough to activate the BIS and prevent participants from cheating. Conversely, the same illusory anonymity can also encourage individuals to point out obvious and fixable embarrassments exhibited by a stranger (e.g., that he has food stuck between his teeth or that his pants zipper is down)—things that are difficult to express but would nonetheless help the stranger to avoid future embarrassment (Bohns et al., 2010).

A heightened cue dependence is similarly observed in anonymous online social behavior, which can produce both supportive self-disclosure and verbal aggression (Joinson, 2007). For instance, Barak and Gluck-Ofri (2007) examined patterns of self-disclosure in online support forums and found that public posts containing higher levels of self-disclosure elicited significantly higher levels of self-disclosure from anonymous commenters. However, these feelings of anonymity also produce greater incidences of negative expressions when the online social environment is less supportive. For example, Moor, Heuvelman, and Verleur (2010) examined the incidence of hostile verbal attacks, or "flaming," on YouTube video comments and found that viewing a greater number of offensive comments posted by others predicted a greater likelihood of contributing similar comments oneself. Being exposed to a greater number of hostile comments led people to perceive a "flaming" norm, thus indicating that this form of expression was context appropriate. This is consistent with the idea that anonymity leads people to follow the most salient response option and to ignore the competing response tendency of social grace.

Anonymity summary

Overall, contexts that produce anonymity tend to result in behaviors with distinct social implications depending on whether the resulting sense of anonymity decreases the perceived social consequences of engaging in selfish and

aggressive behaviors or reduces the potential embarrassment associated with connecting with others through self-disclosure and performing helpful but potentially embarrassing acts. In each of these cases, social concerns produce competing responses when participants are identifiable, resulting in inhibition of the salient action. When anonymous, however, the strength of these social concerns decreases, allowing for the disinhibited expression of the most salient response regardless of its prosocial or antisocial impact.

Implications of a General Model of Disinhibition

A number of novel insights can be obtained by adopting the general model of disinhibition we have articulated in this article. In particular, elaborating upon the central role of the BIS allows multiple domain-specific models of disinhibited behavior to be unified within a common conceptual framework. An additional advantage of our model is that it allows for an integrated perspective on how both situational and dispositional factors can influence the expression of disinhibited behavior, producing both prosocial and antisocial outcomes. Each of the three diverse states we discussed often lead to the same prosocial and antisocial outcomes. For example, increased stereotyping and prejudice is observed as a consequence of power (Fiske, 1993; Goodwin, Gubin, Fiske, & Yzerbyt, 2000; Guinote, Willis, & Martellotta, 2010), alcohol (Bartholow, Dickter, & Sestir, 2006), and anonymity (Lea, Spears, & de Groot, 2001; Plant & Devine, 1998). Conversely, each of these factors can also lead to greater altruism (Bohns et al., 2010; Steele et al., 1985; Whitson et al., 2011).

Disinhibition both reveals and shapes the person

In any given moment, there are a number of potential actions that can be brought to bear on the world. Some of these actions will be externally cued by environmental stimuli, and some will be internally cued by chronically active goals and motivations. We propose that disinhibited states are characterized by a reduced number of potential actions, such that a single dominant behavioral option becomes particularly salient at the expense of other competing options. When an individual is characterized by a strong motivational orientation and does not sense pressing situational cues or norms, the most salient response will be internally driven. However, when the situation is characterized by a narrow range of affordances (i.e., a "strong situation"), the most salient response will be externally driven (Meyer, Dalal, & Hermida, 2010). In contexts where neither the situation nor the disposition is clearly stronger, the most salient responses are likely to be those that receive joint input from both internal and external cues. Understanding the behavioral consequences of disinhibition thus requires an assessment of the most salient responses from both internal (e.g., goals, values, motives, and dispositions) and external (e.g., perceptual cues from the environment) perspectives, the relative strength of each, and the consistency between these internal and external cues (cf. Funder, 2008).

Our model thus explains how disinhibition can both reveal the person, leading to greater correspondence between behavior and underlying dispositions, and shape the person by leading individuals to behave more consistently with strong situational cues. It has been observed, for example, that power can heighten the influence of external cues on behavior (Guinote, 2008), as well as the influence of dispositional characteristics (Galinsky et al., 2008). Previous work has aimed to integrate these two seemingly contradictory perspectives by noting that power can enhance cognitive processing of the primary constructs activated in a situation, whether driven by inner experiences or environmental properties (Guinote, 2007a). We adopt a similar perspective in the current model, emphasizing that power (along with other disinhibiting contexts) can facilitate the engagement of the most salient response, regardless of whether it was triggered by the internal or external environment. A key distinction, however, is that we argue that this is not an effect of power per se, but is a consequence of the disinhibition that power produces. Thus, our model proposes that any disinhibiting context can either reveal or shape the person depending on the relative strength of internal and external response cues. Our model further identifies the reduction of response conflict and BIS activity as the proximal mechanism that underlies such effects (see Figure 1).

Disinhibition produces both prosocial and antisocial outcomes

The model in this article also explains how disinhibition can produce both prosocial and antisocial outcomes. Many models of disinhibition assume that it is universally associated with negative outcomes, such that it involves the expression of socially undesirable behavior and is associated exclusively with psychopathology (Festinger et al., 1952; Nigg, 2000; Zimbardo, 1969). However, our perspective allows both positive and negative outcomes to result from disinhibited states. In Gray's model, behavioral inhibition reflects the slowing or cessation of ongoing behavior, regardless of whether or not that behavior is considered socially desirable (Gray & McNaughton, 2000). Accordingly, there are likely to be many cases, from helping in emergencies to pointing out embarrassing information (e.g., zipper down), in which prosocial behavior is inhibited due to the activation of conflicting responses (e.g., the costs of intervening or wanting to maintain social appropriateness).

We have proposed that reduced BIS activity, driven by the decreased salience of competing response options, is responsible for producing disinhibited behavior. Some recent studies support this notion by demonstrating how directly reducing BIS activity by asking people to recall a time in which they acted without inhibitions can increase both helping behavior and a willingness to cheat for personal benefit (van den Bos,

Müller, & van Bussel, 2009; van den Bos et al., in press). A more balanced perspective on disinhibitory states thus allows for a consideration of both the positive and negative consequences of disinhibition.

The pathways to disinhibition

Although social power, alcohol intoxication, and anonymity appear to be unrelated to one another on the surface, they provide a diverse set of examples for the effects of a common disinhibitory mechanism. These domains also illustrate three different pathways—heightened BAS activity, the depletion of cognitive resources, and the reduction in social desirability concerns—by which the relative strength of any competing motor response is decreased, thereby allowing the most salient action to be expressed without interference from the BIS.

Although we have focused on the specific domains of social power, alcohol intoxication, and anonymity, similar disinhibitory effects are hypothesized to emerge during any context in which the salience of competing response options is decreased and the BIS is correspondingly silenced. Any time the BAS is activated, cognitive resources are depleted, or social desirability concerns are reduced, disinhibition is likely to occur. In addition, these processes could produce either prosocial or antisocial outcomes, depending on the most salient behavioral affordance.

There is emerging evidence in support of the central proposition that any form of BAS activation, reduction in cognitive resources, or decrease in social desirability concerns can result in disinhibition. For example, a dispositionally stronger BAS system has been associated both with a greater tendency to cooperate in a prisoner's dilemma (Hirsh & Peterson, 2009) and a heightened use of physical aggression (Harmon-Jones, 2003). Similarly, the elderly, who often face reductions in cognitive resources, offer more constructive advice in uncomfortable situations (Apfelbaum, Krendl, & Ambady, 2010) but also express greater prejudice (Radvansky, Copeland, & Hippel, 2010). We propose that these apparent contradictions can be resolved by adopting the response salience framework we have outlined. The general model of disinhibition allows for an understanding of when disinhibition is likely to occur, while also predicting whether the social consequences will be positive or negative across a variety of different contexts.

It is worth pointing out, however, that these three distinct pathways to disinhibition can also operate jointly with one another. For example, the powerful are less dependent on others, thereby reducing their concerns about social desirability (Keltner et al., 2003). Similarly, they often face severe demands on their time, leading to attentional overload and depletion (Fiske, 1993). In addition, the depletion of cognitive resources can also lead to a reduced ability to attend to social desirability concerns (e.g., Govorun & Payne, 2006), as well as an increase in BAS activation (Schmeichel, Harmon-Jones, & Harmon-Jones, 2010). More generally, it is important to consider the multiple disinhibiting forces that might be present in any given context, as these pathways may be operating in parallel and be mutually reinforcing.

Future directions

A number of novel empirical predictions also emerge from our proposal that decreases in response conflict and the corresponding reductions in BIS activity serve as a common mechanism that produces behavioral disinhibition. First, if the diverse contexts of inebriation, power, and anonymity do indeed exert their behavioral effects through a common mechanism, then a variety of interactions should be observed when the manipulations are administered simultaneously. One possibility is that disinhibition works according to a threshold effect (Inesi, Botti, Dubois, Rucker, & Galinsky, in press); once one source or pathway has produced a disinhibited state (e.g., power), the addition of another source of disinhibition (e.g., alcohol) will not produce more disinhibited behavior. For example, alcohol intoxication might reduce the behavioral differences between the powerful and the powerless. Similarly, a state of anonymity could reduce the behavioral differences that are observed between sober and intoxicated individuals. In each of these cases, the individual will already be in a disinhibited state, thereby reducing the behavioral impact of any further disinhibiting force. Alternatively, the sources could combine in additive ways. For example, the anonymous drunks might be particularly disinhibited. Future research should explore whether the sources of and pathways to disinhibition are additive or operate according to a threshold effect.

Second, the central role of the BIS in our framework allows for novel predictions to be made from an individual differences perspective, given that the dispositional strength of the BIS is thought to be reflected in the personality trait of neuroticism (Gray & McNaughton, 2000). Higher levels of neuroticism are associated with a more sensitive BIS, such that neurotic individuals have a lower threshold for behavioral inhibition in response to potentially conflicting actions. Consequently, two novel yet opposing hypotheses emerge from our model. First, given their higher baseline levels of BIS activity, the behavioral difference between inhibited and disinhibited states may be most pronounced among neurotic individuals. Highly anxious individuals would thus demonstrate the largest observable change in their behavior when drunk rather than sober, when they are powerful rather than powerless, or when they are anonymous rather than identifiable. In contrast, it may be the case that more neurotic individuals retain their relatively higher levels of BIS activation even when competing responses are less salient. Consequently, the effects of these disinhibiting situations may be less pronounced at higher levels of neuroticism, such that neurotic individuals would remain inhibited. Although both of these effects are plausible, available evidence appears to support the first hypothesis: Neurotic individuals are more likely to seek out alcohol and anonymity as a means of reducing the experience of anxiety and the concomitant BIS activation (Kuntsche,

Knibbe, Gmel, & Engels, 2006; Shepherd & Edelmann, 2005). Future research can help to address this question more directly.

Third, identifying the BIS as a common mechanism in each of these cases allows specific neurophysiological predictions to be made. Because our framework argues that reduced response conflict is the key factor underlying disinhibited behavior, neural indicators of response conflict should be reduced in disinhibited situations. For instance, the Error-Related Negativity (ERN) is an event-related potential component that is implicated in the detection of conflict, performance errors, and uncertainty (Yeung et al., 2004). It has also been related to the BIS, such that heightened BIS activity tends to be associated with a larger ERN (Amodio et al., 2008; Boksem et al., 2006). If the disinhibiting contexts that we described do indeed exert their effects through a reduction in response conflict and the corresponding decreases in BIS activity, individuals in these disinhibited states would be expected to demonstrate a smaller ERN. Initial evidence is supportive of this possibility, as alcohol intoxication is indeed associated with smaller ERNs in response to performance errors (Curtin & Fairchild, 2003; Ridderinkhof et al., 2002; Yeung & Cohen, 2006). Future studies should examine whether the same holds for power and anonymity.

Summary

Social power, alcohol intoxication, and anonymity can all produce seemingly paradoxical behavioral outcomes. In some cases, they can increase prosocial behavior, whereas in others they promote antisocial or selfish behavior. The contradictory effects of these diverse manipulations can be explained by pointing to a common mechanism of disinhibition. Thus, our general model of disinhibition has attempted to unify multiple domain-specific models of disinhibited behavior into a common conceptual framework. Although the specific routes to disinhibition are diverse (activation of the BAS, cognitive resource depletion, and reduction in social desirability concerns), they all involve the reduced activation of competing behavioral options, thereby preventing BIS activation and ultimately producing the uninhibited expression of the most salient response. The specific response that emerges depends on both preexisting dispositional motives (strong internal cues) and the dominant affordances of the situation (strong external cues). Thus, disinhibition can both reveal and shape the person. Taking into account an individual's dispositional orientation and the situational context allows for a better understanding of the behavioral consequences of disinhibited states. Although the powerful and inebriated might be more willing to help during an emergency, it may be best to turn the lights on and hide the alcohol when relying on them to avoid their selfish inclinations.

References

Amodio, D. M., Master, S. L., Yee, C. M., & Taylor, S. E. (2008). Neurocognitive components of the behavioral inhibition and activation systems: Implications for theories of self-regulation. *Psychophysiology*, *45*, 11–19.

Anderson, C., & Berdahl, J. L. (2002). The experience of power: Examining the effects of power on approach and inhibition tendencies. *Journal of Personality and Social Psychology*, *83*, 1362–1377.

Apfelbaum, E. P., Krendl, A. C., & Ambady, N. (2010). Age-related decline in executive function predicts better advice-giving in uncomfortable social contexts. *Journal of Experimental Social Psychology*, *46*, 1074–1077.

Ariely, D., & Loewenstein, G. (2006). The heat of the moment: The effect of sexual arousal on sexual decision making. *Journal of Behavioral Decision Making*, *19*, 87–98.

Awh, E., Vogel, E. K., & Oh, S. H. (2006). Interactions between attention and working memory. *Neuroscience*, *139*, 201–208.

Baddeley, A. (1986). *Working memory*. Oxford, England: Oxford University Press.

Barak, A., & Gluck-Ofri, O. (2007). Degree and reciprocity of self-disclosure in online forums. *CyberPsychology and Behavior*, *10*, 407–417.

Bargh, J. A., McKenna, K. Y. A., & Fitzsimons, G. M. (2002). Can you see the real me? Activation and expression of the "true self" on the Internet. *Journal of Social Issues*, *58*, 33–48.

Bargh, J. A., Raymond, P., Pryor, J. B., & Strack, F. (1995). Attractiveness of the underling: An automatic power sex association and its consequences for sexual harassment and aggression. *Journal of Personality and Social Psychology*, *68*, 768–781.

Bartholow, B. D., Dickter, C. L., & Sestir, M. A. (2006). Stereotype activation and control of race bias: Cognitive control of inhibition and its impairment by alcohol. *Journal of Personality and Social Psychology*, *90*, 272–287.

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

Bohns, V. K., Gino, F., & Zhong, C. B. (2010). *A brighter side to darkness: The illusion of anonymity can increase prosocial behavior*. Unpublished Manuscript.

Boksem, M. A. S., Smolders, R., & De Cremer, D. (2009). Social power and approach-related neural activity. *Social Cognitive and Affective Neuroscience*. Epub ahead of print.

Boksem, M. A. S., Tops, M., Wester, A. E., Meijman, T. F., & Lorist, M. M. (2006). Error-related ERP components and individual differences in punishment and reward sensitivity. *Brain Research*, *1101*, 92–101.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J.D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*, 624–652.

Cannavale, F. J., Scarr, H. A., & Pepitone, A. (1970). Deindividuation in the small group: Further evidence. *Journal of Personality and Social Psychology*, *16*, 141–147.

Carver, C. S., Sutton, S. K., & Scheier, M. F. (2000). Action, emotion, and personality: Emerging conceptual integration. *Personality and Social Psychology Bulletin*, *26*, 741–751.

Carver, C. S ., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology*, *67*, 319–333.

Chen, S., Lee-Chai, A. Y., & Bargh, J. A. (2001). Relationship orientation as a moderator of the effects of social power. *Journal of Personality and Social Psychology*, *80*, 173–187.

Cooper, M. L. (2006). Does drinking promote risky sexual behavior? *Current Directions in Psychological Science*, *15*, 19–23.

Corr, P. J. (2002). J. A. Gray's reinforcement sensitivity theory: Tests of the joint subsystems hypothesis of anxiety and impulsivity. *Personality and Individual Differences*, *33*, 511–532.

Corr, P. J. (2004). Reinforcement sensitivity theory and personality. *Neuroscience & Biobehavioral Reviews*, *28*, 317–332.

Côté, S., Kraus, M. W., Cheng, B. H., Oveis, C., van der Löwe, I., Lian, H., & Keltner, D. (2011). Social power facilitates the effect of prosocial orientation on empathic accuracy. *Journal of Personality and Social Psychology*, *101*, 217–232.

Curtin, J. J., & Fairchild, B.A. (2003). Alcohol and cognitive control: Implications for regulation of behavior during response conflict. *Journal of Abnormal Psychology*, *112*, 424–436.

Davis, K. C., Hendershot, C. S., & George, W. H. (2007). Alcohol's effects on sexual decision making: An integration of alcohol myopia and individual differences. *Alcohol*, *68*, 843–851.

Depue, R. A., & Collins, P. F. (1999). Neurobiology of the structure of personality: Dopamine, facilitation of incentive motivation, and extraversion. *Behavioral and Brain Sciences*, *22*, 491–517.

Diener, E. (1979). Deindividuation, self-awareness, and disinhibition. *Journal of Personality and Social Psychology*, *37*, 1160–1171.

Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation*. Hillsdale, NJ: Erlbaum.

Eagly, A. H., & Steffen, V. J. (1986). Gender and aggressive behavior: A meta-analytic review of the social psychological literature. *Psychological Bulletin*, *100*, 309–330.

Easdon, C. M., & Vogel-Sprott, M. (2000). Alcohol and behavioral control: Impaired response inhibition and flexibility in social drinkers. *Experimental and Clinical Psychopharmacology*, *8*, 387–394.

Elliot, A. J., & Thrash, T. M. (2002). Approach-avoidance motivation in personality: Approach and avoidance temperaments and goals. *Journal of Personality and Social Psychology*, *82*, 804–818.

Emerson, R. M. (1962). Power-dependence relations. *American Sociological Review*, *27*, 31–41.

Emmons, R. A., & King, L. A. (1988). Conflict among personal strivings: Immediate and long-term implications for psychological and physical well-being. *Journal of Personality and Social Psychology*, *54*, 1040–1048.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*, 19–23.

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102–134). New York, NY: Cambridge University Press.

Festinger, L., Pepitone, A., & Newcomb, T. (1952). Some consequences of de-individuation in a group. *The Journal of Abnormal and Social Psychology*, *47*, 382–389.

Fillmore, M. T., & Vogel-Sprott, M. (1999). An alcohol model of impaired inhibitory control and its treatment in humans. *Experimental and Clinical Psychopharmacology*, *7*, 49–55.

Fiske, S. T. (1993). Controlling other people. *American Psychologist*, *48*, 621–628.

Frone, M. R. (2000). Work-family conflict and employee psychiatric disorders: The National Comorbidity Survey. *Journal of Applied Psychology*, *85*, 888–896.

Funder, D. C. (2008). Persons, situations, and person-situation interactions. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 568–580). New York, NY: Guilford Press.

Gable, P., & Harmon-Jones, E. (2008). Approach-motivated positive affect reduces breadth of attention. *Psychological Science*, *19*, 476–482.

Gable, P., & Harmon-Jones, E. (2010). The motivational dimensional model of affect: Implications for breadth of attention, memory, and cognitive categorisation. *Cognition & Emotion*, *24*, 322–337.

Galinsky, A. D., Gruenfeld, D. H., & Magee, J. C. (2003). From power to action. *Journal of Personality and Social Psychology*, *85*, 453–466.

Galinsky, A. D., Magee, J. C., Gruenfeld, D. H., Whitson, J. A., & Liljenquist, K.A. (2008). Power reduces the press of the situation: Implications for creativity, conformity, and dissonance. *Journal of Personality and Social Psychology*, *95*, 1450–1466.

Galinsky, A. D., Rus, D., & Lammers, J. (2010). Power: A central force governing psychological, social, and organizational life. In D. De Cremer, J. K. Murnighan, & R. van Dick (Eds.), *Social psychology and organizations* (pp. 17–38). Philadelphia, PA: Psychology Press.

Gergen, K. J., Gergen, M. M., & Barton, W. H. (1973). Deviance in the dark. *Psychology Today*, *7*, 129–133.

Giancola, P. R. (2002). Alcohol-related aggression in men and women: The influence of dispositional aggressivity. *Journal of Studies on Alcohol*, *63*, 696–709.

Giancola, P. R. (2003). The moderating effects of dispositional empathy on alcohol-related aggression in men and women. *Journal of Abnormal Psychology*, *112*, 275–281.

Goodwin, S. A., Gubin, A., Fiske, S. T., & Yzerbyt, V. Y. (2000). Power can bias impression processes: Stereotyping subordinates by default and by design. *Group Processes & Intergroup Relations*, *3*, 227–256.

Govorun, O., & Payne, B. K. (2006). Ego-depletion and prejudice: Separating automatic and controlled components. *Social Cognition*, *24*, 111–136.

Gray, J. A. (1982). *The neuropsychology of anxiety*. New York, NY: Oxford University Press.

Gray, J. A. (1990). Brain systems that mediate both emotion and cognition. *Cognition & Emotion*, *4*, 269–288.

Gray, J. A., & McNaughton, N. (2000). *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system*. Oxford, England: Oxford University Press.

Guinote, A. (2007a). Behaviour variability and the situated focus theory of power. *European Review of Social Psychology*, *18*, 256–295.

Guinote, A. (2007b). Power and goal pursuit. *Personality and Social Psychology Bulletin*, *33*, 1076–1087.

Guinote, A. (2008). Power and affordances: When the situation has more power over powerful than powerless individuals. *Journal of Personality and Social Psychology*, *95*, 237–252.

Guinote, A., & Vescio, T. K. (Eds.). (2010). *The social psychology of power*. New York, NY: Guilford Press.

Guinote, A., Willis, G. B., & Martellotta, C. (2010). Social power increases implicit prejudice. *Journal of Experimental Social Psychology*, *46*, 299–307.

Gustafson, R. (1993). What do experimental paradigms tell us about alcohol-related aggressive responding? *Journal of Studies on Alcohol*, *11*(Suppl.), 20–29.

Hajcak, G., McDonald, N., & Simons, R. F. (2003). Anxiety and error-related brain activity. *Biological Psychology*, *64*, 77–90.

Handgraaf, M. J. J., Van Dijk, E., Vermunt, R. C., Wilke, H. A. M., & De Dreu, C. K. W. (2008). Less power or powerless? Egocentric empathy gaps and the irony of having little versus no power in social decision making. *Journal of Personality and Social Psychology*, *95*, 1136–1149.

Harmon-Jones, E. (2003). Anger and the behavioral approach system. *Personality and Individual Differences*, *35*, 995–1005.

Harmon-Jones, E., & Harmon-Jones, C. (2002). Testing the action-based model of cognitive dissonance: The effect of action orientation on postdecisional attitudes. *Personality and Social Psychology Bulletin*, *28*, 711–723.

Hirsh, J. B., & Peterson, J. B. (2009). Extraversion, neuroticism, and the prisoner's dilemma. *Personality and Individual Differences*, *46*, 254–256.

Inesi, M. E. (2010). Power and loss aversion. *Organizational Behavior and Human Decision Processes*, *112*, 58–69.

Inesi, M. E., Botti, S., Dubois, D., Rucker, D. D., & Galinsky, A. D. (in press). Power and choice: Their dynamic interplay in quenching the thirst for personal control. *Psychological Science*.

Janssen, E., Vorst, H., Finn, P., & Bancroft, J. (2002). The Sexual Inhibition (SIS) and Sexual Excitation (SES) Scales: I. Measuring sexual inhibition and excitation proneness in men. *Journal of Sex Research*, *39*, 114–126.

Jeavons, C. M., & Taylor, S. P. (1985). The control of alcohol-related aggression: Redirecting the inebriate's attention to socially appropriate conduct. *Aggressive Behavior*, *11*, 93–101.

Johnson, R. D., & Downing, L. L. (1979). Deindividuation and valence of cues: Effects on prosocial and antisocial behavior. *Journal of Personality and Social Psychology*, *37*, 1532–1538.

Joinson, A. N. (1999). Anonymity, disinhibition and social desirability on the Internet. *Behaviour Research Methods, Instruments and Computers*, *31*, 433–438.

Joinson, A. N. (2007). Causes and implications of disinhibited behavior on the Internet. In J. Gackenbach (Ed.), *Psychology and the Internet* (pp. 43–60). Boston, MA: Academic Press.

Keltner, D., Gruenfeld, D. H., & Anderson, C. (2003). Power, approach, and inhibition. *Psychological Review*, *110*, 265–284.

Kuntsche, E., Knibbe, R., Gmel, G., & Engels, R. (2006). Who drinks and why? A review of socio-demographic, personality, and contextual issues behind the drinking motives in young people. *Addictive Behaviors*, *31*, 1844–1857.

Lammers, J., Stapel, D. A., & Galinsky, A. D. (2010). Power increases hypocrisy. *Psychological Science*, *21*, 737–744.

Lea, M., Spears, R., & de Groot, D. (2001). Knowing me, knowing you: Anonymity effects on social identity processes within groups. *Personality and Social Psychology Bulletin*, *27*, 526–537.

Leary, M. R., & Kowalski, R. M. (1990). Impression management: A literature review and two-component model. *Psychological Bulletin*, *107*, 34–47.

Luu, P., Tucker, D. M., Derryberry, D., Reed, M., & Poulsen, C. (2003). Electrophysiological responses to errors and feedback in the process of action regulation. *Psychological Science*, *14*, 47–53.

MacDonald, A. W., Cohen, J. D., Stenger, V.A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*, *288*, 1835–1838.

MacDonald, T. K., Fong, G. T., Zanna, M. P., & Martineau, A. M. (2000). Alcohol myopia and condom use: Can alcohol intoxication be associated with more prudent behavior? *Journal of Personality*, *78*, 605–619.

Magee, J. C., & Galinsky, A. D. (2008). Social hierarchy: The self-reinforcing nature of power and status. *Academy of Management Annals*, *2*, 351–398.

Mast, M. S., Jonas, K., & Hall, J. A. (2009). Give a person power and he or she will show interpersonal sensitivity: The phenomenon and its why and when. *Journal of Personality and Social Psychology*, *97*, 835–850.

McGregor, I., Nash, K., Mann, N., & Phills, C. E. (2010). Anxious uncertainty and reactive approach motivation (RAM). *Journal of Personality and Social Psychology*, *99*, 133–147.

McNaughton, N., & Gray, J. A. (2000). Anxiolytic action on the behavioural inhibition system implies multiple types of arousal contribute to anxiety. *Journal of Affective Disorders*, *61*, 161–176.

Meyer, R. D., Dalal, R. S., & Hermida, R. (2010). A review and synthesis of situational strength in the organizational sciences. *Journal of Management*, *36*, 121–140.

Moeller, F. G., Dougherty, D. M., Lane, S. D., Steinberg, J. L., & Cherek, D. R. (1998). Antisocial personality disorder and alcohol-induced aggression. *Alcoholism: Clinical and Experimental Research*, *22*, 1898–1902.

Moor, P. J., Heuvelman, A., & Verleur, R. (2010). Flaming on YouTube. *Computers in Human Behavior*, *26*, 1536–1546.

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin*, *126*, 220–246.

Olvet, D. M., & Hajcak, G. (2008). The error-related negativity (ERN) and psychopathology: Toward an endophenotype. *Clinical Psychology Review*, *28*, 1343–1354.

Patterson, C. M., & Newman, J. P. (1993). Reflectivity and learning from aversive events: Toward a psychological mechanism for the syndromes of disinhibition. *Psychological Review*, *100*, 716–736.

Peterson, J. B., Pihl, R. O., Seguin, J.R., Finn, P. R., & Stewart, S.H. (1993). Alcohol-induced heart rate change, family history, and prediction of weekly alcohol consumption by nonalcoholic males. *Journal of Studies on Alcohol*, *53*, 499–506.

Pihl, R. O., Peterson, J. B., & Lau, M. A. (1993). A biosocial model of the alcohol-aggression relationship. *Journal of Studies on Alcohol and Drugs*, *11*, 128–139.

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology*, *75*, 811–832.

Postmes, T., & Spears, R. (1998). Deindividuation and antinormative behavior: A meta-analysis. *Psychological Bulletin*, *123*, 238–259.

Prentice-Dunn, S., & Rogers, R. W. (1980). Effects of deindividuating situational cues and aggressive models on subjective deindividuation and aggression. *Journal of Personality and Social Psychology*, *39*, 104–113.

Radvansky, G. A., Copeland, D. E., & Hippel, W. (2010). Stereotype activation, inhibition, and aging. *Journal of Experimental Social Psychology*, *46*, 51–60.

Reicher, S. (2001). The psychology of crowd dynamics. In M. A. Hogg & R. S. Tindale (Eds.), *Blackwell handbook of social psychology: Group processes* (pp. 182–208). Hoboken, NJ: Blackwell.

Ridderinkhof, K. R., de Vlugt, Y., Bramlage, A., Spaan, M., Elton, M., Snel, J., & Band, G. P. H. (2002). Alcohol consumption impairs detection of performance errors in mediofrontal cortex. *Science*, *298*, 2209.

Russell, G. W. (1995). Personalities in the crowd: Those who would escalate a sports riot. *Aggressive Behavior*, *21*, 91–100.

Russell, G. W. (2004). Sport riots: A social-psychological review. *Aggression and Violent Behavior*, *9*, 353–378.

Russell, G. W., & Mustonen, A. (1998). Peacemakers: Those who would intervene to quell a sports riot. *Personality and Individual Differences*, *24*, 335–339.

Schmeichel, B. J., Harmon-Jones, C., & Harmon-Jones, E. (2010). Exercising self-control increases approach motivation. *Journal of Personality and Social Psychology*, *99*, 162–173.

Shah, J.Y., Friedman, R., & Kruglanski, A.W. (2002). Forgetting all else: On the antecedents and consequences of goal shielding. *Journal of Personality and Social Psychology*, *83*, 1261–1280.

Shepherd, R. M., & Edelmann, R. J. (2005). Reasons for Internet use and social anxiety. *Personality and Individual Differences*, *39*, 949–958.

Smith, P. K., & Bargh, J. A. (2008). Nonconscious effects of power on basic approach and avoidance tendencies. *Social Cognition*, *26*, 1–24.

Steele, C. M., Critchlow, B., & Liu, T. J. (1985). Alcohol and social behavior II: The helpful drunkard. *Journal of Personality and Social Psychology*, *48*, 35–46.

Steele, C. M., & Josephs, R. A. (1990). Alcohol myopia: Its prized and dangerous effects. *American Psychologist*, *45*, 921–933.

Steele, C. M., & Southwick, L. (1985). Alcohol and social behavior 1: The psychology of drunken excess. *Journal of Personality and Social Psychology*, *48*, 18–34.

Suler, J. (2004). The online disinhibition effect. *CyberPsychology & Behavior*, *7*, 321–326.

Sutton, S. K., & Davidson, R. J. (1997). Prefrontal brain asymmetry: A biological substrate of the behavioral approach and inhibition systems. *Psychological Science*, *8*, 204–210.

Tjosvold, D. (1985). Power and social context in superior-subordinate interaction. *Organizational Behavior and Human Decision Processes*, *35*, 281–293.

van den Bos, K., Müller, P. A., & van Bussel, A. A. L. (2009). Helping to overcome intervention inertia in bystander's dilemmas: Behavioral disinhibition can improve the greater good. *Journal of Experimental Social Psychology*, *45*, 873–878.

van den Bos, K., Van Lange, P.A.M., Lind, E. A., Venhoeven, L. A., Beudeker, D. A., & Cramwinckel, F. M., . . . Van der Laan, J. (in press). On the benign qualities of behavioral disinhibition: Because of the prosocial nature of people behavioral disinhibition can weaken pleasure with getting more than you deserve. *Journal of Personality and Social Psychology*.

Vohs, K. D., & Heatherton, T. F. (2000). Self-regulatory failure: A resource-depletion approach. *Psychological Science*, *11*, 249–254.

von Hippel, W. (2007). Aging, executive functioning, and social control. *Current Directions in Psychological Science*, *16*, 240–244.

Ward, A., & Mann, T. (2000). Don't mind if I do: Disinhibited eating under cognitive load. *Journal of Personality and Social Psychology*, *78*, 753–763.

Watson, D., Wiese, D., Vaidya, J., & Tellegen, A. (1999). The two general activation systems of affect: Structural findings, evolutionary considerations and psychobiological evidence. *Journal of Personality and Social Psychology*, *76*, 820–838.

Whitson, J. A., Liljenquist, K. A., Galinsky, A. D., Magee, J. C., & Gruenfeld, D. H. (2011). *Powerful saviors: Priming power increases bystander intervention in emergencies*. Unpublished Manuscript.

Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error detection: Conflict monitoring and the error-related negativity. *Psychological Review*, *111*, 931–959.

Yeung, N., & Cohen, J.D. (2006). The impact of cognitive deficits on conflict monitoring. *Psychological Science*, *17*, 164–171.

Zhong, C. B., Bohns, V. K., & Gino, F. (2010). Good lamps are the best police: Darkness increases dishonesty and self-interested behavior. *Psychological Science*, *21*, 311–314.

Zhong, C. B., Galinsky, A. D., Magee, J. C., & Maddux, W. W. (2011). *Culture moderates the automatic associations with power*. Unpublished Manuscript.

Zimbardo, P. G. (1969). The human choice: Individuation, reason, and order versus deindividuation, impulse, and chaos. In W. J. Arnold & D. Levine (Eds.), *Nebraska symposium on motivation* (Vol. 17, pp. 237–307). Lincoln: University of Nebraska Press.