A Two-Dimensional Model for the Study of Interpersonal Attraction
R. Matthew Montoya and Robert S. Horton
Pers Soc Psychol Rev 2014 18: 59 originally published online 10 September 2013
DOI: 10.1177/1088868313501887

The online version of this article can be found at:
http://psr.sagepub.com/content/18/1/59

Published by:
SAGE
http://www.sagepublications.com

On behalf of:

Society for Personality and Social Psychology

Additional services and information for Personality and Social Psychology Review can be found at:

Email Alerts: http://psr.sagepub.com/cgi/alerts

Subscriptions: http://psr.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

>> Version of Record - Jan 16, 2014

OnlineFirst Version of Record - Sep 10, 2013

What is This?
A Two-Dimensional Model for the Study of Interpersonal Attraction

R. Matthew Montoya¹ and Robert S. Horton²

Abstract
We describe a model for understanding interpersonal attraction in which attraction can be understood as a product of the initial evaluations we make about others. The model posits that targets are evaluated on two basic dimensions, capacity and willingness, such that affective and behavioral attraction result from evaluations of (a) a target’s capacity to facilitate the perceiver’s goals/needs and (b) a target’s potential willingness to facilitate those goals/needs. The plausibility of the two-dimensional model of attraction is evaluated vis-à-vis the extant literature on various attraction phenomena including the reciprocity of liking effect, pratfall effect, matching hypothesis, arousal effects, and similarity effect. We conclude that considerable evidence across a wide range of phenomena supports the idea that interpersonal attraction is principally determined by inferences about the target’s capacity and willingness.

Keywords
interpersonal processes, person perception, attitudes

The field has seen tremendous methodological and theoretical advances in the study of interpersonal attraction over the past 75 years. Since Moreno (1934) developed the first rigorous experimental investigation of interpersonal attraction between individuals (termed the “sociometric method”), researchers have explored attraction in the laboratory and the field, in real and artificial relationships, and in the context of romantic and non-romantic dyads. Despite the numerous advancements and discoveries, serious critiques regarding attraction research have persisted. For instance, researchers have disagreed about (a) how to measure or conceptualize attraction (e.g., Griffin & Langlois, 2006; Huston, 1974; D. Marlowe & Gergen, 1969; Rosenbaum, 1986), (b) the theoretical orientation to use to study attraction (Lefkowitz, 2000), (c) the need to study attraction across time (i.e., relationship development; Kenny, 1994; Reis, 2007), and (d) the adequacy of experimental work for elucidating the nature and course of interpersonal attraction (Berscheid, 1985; Berscheid & Graziano, 1979; Duck & Allison, 1978; Graziano & Bruce, 2008; Reis, 2007). The consequence of these disagreements has been a fragmented field, one that proceeds without a unifying theoretical model that can synthesize the impressive corpus of research that has accumulated over the years.

The goal of this article is to address the disagreements noted above and present a unified model of interpersonal attraction. In so doing, we (a) discuss the definition and measurement of the attraction construct, (b) propose a two-dimensional model of attraction based on person perception and relationship research, and (c) describe how such a model can account for a wide range of seemingly disparate findings related to interpersonal attraction. Specifically, this model proposes that two dimensions fundamental to person perception, capacity and willingness, influence interpersonal attraction by indicating a target’s (a) capacity to facilitate the needs/goals of a perceiver and (b) willingness to facilitate such needs and goals, respectively. That is, affective and behavioral attraction can be understood as a consequence of assessments that indicate a target’s willingness and capacity to facilitate the individual’s goals and interests. Before outlining the model and evaluating it in the context of established attraction phenomena, we first turn to the definition of the interpersonal attraction construct.

Definition of Interpersonal Attraction
Among the primary obstacles to synthesizing the attraction literature are confusion and disagreement about the definition of “atraction.” Interpersonal attraction is commonly thought of as a positive emotional evaluation of another person (i.e., an affective evaluation), but it is often operationalized behaviorally (e.g., degree of affiliation, proximity of chair placement) or cognitively (e.g., assessment of traits/attributes). The term, attraction, originated in the 1500s as a

¹University of Dayton, OH, USA
²Wabash College, Crawfordsville, IN, USA

Corresponding Author:
R. Matthew Montoya, University of Dayton, 300 College Park, Dayton, OH 45469, USA.
Email: matt.montoya@udayton.edu
medical term and referred to the body’s tendency to absorb fluids or nourishment (Oxford English Dictionary, 2013). Over the subsequent centuries, it evolved to refer to the ability for an object to draw an object to itself, and then to the ability of a person to draw another person to him or her. Whereas some researchers have maintained this focus on behavior (i.e., “drawing one to another,” for example, Schachter, 1959), others have emphasized the affective (i.e., feeling positively toward; Zajonc, 1968) or cognitive (i.e., inferring positive attributes; R. Singh, Ho, Tan, & Bell, 2007) dimensions. The consistent aspect of these definitions is the description of a positive evaluation of another person (Berscheid, 1985; Huston & Levinger, 1978). As a construct requiring measurement, attraction necessitates a discrete definition that acknowledges its varied qualities.

Whereas past reviews of the attraction literature have used the tripartite model (e.g., Berscheid, 1985) to define attraction (i.e., attraction has an affective, behavioral, and cognitive component; Katz & Stotland, 1959; M. Rosenberg, Hovland, McGuire, Abelson, & Brehm, 1960), such a synthetic definition ignores important distinctions and causal relations among the components. We define attraction as a person’s immediate and positive affective and/or behavioral response to a specific individual, a response that is influenced by the person’s cognitive assessments. By this thinking, attraction has two valenced components—an affective component that reflects the quality of one’s emotional response toward another (commonly measured as physiological responses or self-reports of the emotional response) and a behavioral component that reflects one’s tendency to act in a particular way toward another (often measured as choices to affiliate or as actual/preferred proximity). For simplicity, we use the term “attraction” or “liking” when the individual components (affective and behavioral) are consonant; however, in some cases, the behavioral and affective components of attraction can diverge. In such cases, we refer to “behavioral attraction” and “affective attraction” to describe the two components.

With respect to the cognitive component, we argue that attraction per se does not include a cognitive component but rather is preceded and influenced by two distinct cognitive assessments of a target: a capacity assessment and a willingness assessment, which we will describe in detail. There is considerable evidence to support excluding the cognitive component from the operationalization of attraction. First, cognitive responses have often been assessed as an outcome that is different from attraction phenomenon. As an example, research on the physical attractiveness stereotype characterizes beliefs about a target’s sociability or integrity as separate from whether one is attracted to that target (e.g., Eagly, Ashmore, Makihjani, & Longo, 1991). Second, cognitive responses are often measured and modeled as a precursor to attraction. Information integration models, for example, treat cognitive assessments as a mediator of attraction (e.g., Kaplan & Anderson, 1973). Third, traditional operationalizations of attraction have emphasized and discussed the affective and behavioral assessments but not cognitive ones (e.g., Berscheid, 1985; Huston & Levinger, 1978). Fourth, some attraction phenomena (e.g., mere exposure effect, Zajonc, 1980; implicit egotism, Pelham, Carvallo, & Jones, 2005) are theorized to occur in the absence of cognitive responses. Finally, as will be discussed, research has noted that attitude formation during the person perception process involves two distinct cognitive evaluations (e.g., Peeters, 2001; Todorov, 2011; Wojciszke, 2005), indicating that the nature of person perception may be more complex than the cognitive component proposed by the classic tripartite model.

The relation between affective and behavioral attraction also requires further discussion. These two components of attraction are likely to be strongly related in most circumstances. However, because the two components may be influenced by different factors, a discussion of such divergence is critical to an understanding of the attraction process. To begin, researchers (e.g., D. Davis, 1981; Werner & Parmelee, 1979) posit that behavioral evaluations are more dependent on social context and self-interest considerations than are affective evaluations. Such dependence on context and self-interest results from the importance of behavioral attraction to increasing (a) the likelihood of receiving potential rewards from an interaction (Michinov & Monteil, 2002) and (b) the amount of information gained about the other person (D. Davis, 1981). Thus, whether a target can benefit or undermine one’s self-interests should be reflected more directly in behavioral attraction than in affective attraction. As examples, imagine that you had an acquaintance whom you did not necessarily like but who had a car and who had offered you a ride to work, an offer that your car-less self would benefit from accepting. In this case, affective attraction would be rather low (you do not feel positively toward the acquaintance), but behavioral attraction might be rather high (you would go out of your way to affiliate with the acquaintance). Conversely, one might also imagine experiencing affective, but not behavioral, attraction to a target. For example, a person might experience intense affective attraction to a supermodel at the local discotheque, but experience little to no behavioral attraction due to the high probability of romantic rejection. In the latter case, behavioral attraction is more sensitive to self-interests (e.g., avoidance of rejection) than is affective attraction.

As defined here, attraction describes one’s instantized evaluation of many different possible people, including sexual or romantic partners, friends, co-workers, neighbors, or unmet others. At the same time, attraction should not be confused with a number of other relationship-oriented constructs, including romantic or passionate love, friendship, attachment, kinship, and limerence. Each of these related concepts describes a relatively enduring interpersonal experience (as compared with attraction), rather than an immediate evaluation of a target person, and/or describes an experience that is specific to a particular type of relationship...
(e.g., a potential sexual, non-sexual, or family relationship) rather than one that characterizes interpersonal experiences generally, as attraction does. Furthermore, attraction should not be regarded as simply the “opposite” of negatively valued constructs such as repulsion, dislike, or prejudice (for a discussion, see Cacioppo, Gardner, & Berntson, 1997; Gable, 2006; Pittinsky, Rosenthal, & Montoya, 2011; Winkielman & Cacioppo, 2001).

**Two-Dimensional Models in Relationship and Person Perception Research**

An understanding of the interpersonal attraction process begins with how humans perceive and initially evaluate others. The person perception and relationship literature has provided tremendous consensus regarding the two fundamental dimensions on which humans make social judgments: one dimension that assesses the capacity (competence/skill) of the other person and a second that assesses the willingness (trustworthiness/warmth) of the other person to benefit the perceiver’s goals/interests (Fiske, 1992; Wojciszke, 2005). In the following section, we review the person perception findings and then review research that has applied these dimensions to the relationship literature.

**Models of Person Perception**

Early person perception research identified warm–cold and competence-related adjectives (Asch, 1946) as critical to interpersonal evaluations. Additional research confirmed these two dimensions using multidimensional scaling (S. Rosenberg, Nelson, & Vivekananthan, 1968; see also Lydon, Jamieson, & Zanna, 1988) and anthropological evidence pointed toward these same two dimensions (White, 1980). Recent empirical work, across multiple laboratories and in a variety of research contexts, has further emphasized the importance of these two dimensions for interpersonal judgment.

Specifically, multiple contemporary laboratories investigating how people evaluate persons, groups, and faces have each identified analogous dimensions and have noted that these dimensions principally comprise one’s evaluations of people and groups. Wojciszke and colleagues (Wojciszke, 1994, 2005; Wojciszke & Klusek, 1996), for example, posited that the two fundamental dimensions of person perception are morality and competence. According to Wojciszke (2005), morality traits (such as honest, trustworthy, egocentric) refer to goals of the person and the relation of those goals to moral norms. Competency traits (such as intelligent, capable, industrious) refer to the efficiency with which those moral goals are attained. Alternatively, Peeters and colleagues (e.g., Peeters, 1983, 2001; Peeters & Czapinski, 1990) described self-profitability and other-profitability as the two critical dimensions of person perception. Self-profitability refers to traits that reflect the possessor’s chances of achieving personal goals and includes traits such as competence, intelligence, and efficiency. Other-profitability refers to traits that affect the outcomes of other people, rather than the possessor, and includes traits such as generous, kind, selfish, and trustworthy. When participants were asked to categorize traits by the degree to which they were associated with morality, competence, and other- versus self-profit, other- and self-profitability accounted for 97% of the variance of global evaluations (Wojciszke, Dowhyluk, & Jaworski, 1998).

Research has also confirmed these two dimensions as fundamental to the evaluation of social groups (e.g., Clausell & Fiske, 2005; Cuddy, Fiske, & Glick, 2004; Glick, 2002). Fiske and colleagues (e.g., Fiske, Cuddy, Glick, & Xu, 2002; Fiske, Xu, Cuddy, & Glick, 1999) described these dimensions as competence (traits related to the other’s groups position in a status hierarchy [e.g., skill]) and warmth (traits related to the group’s intention to cooperate/compete with other groups [e.g., morality]). It is noteworthy that these dimensions, which were derived to describe evaluations of social groups, have been described by these researchers as analogs to those dimensions identified at the individual level by Wojciszke and colleagues and Peeters and colleagues (Fiske et al., 2002).

Finally, research on facial evaluations also conforms to the two-dimensional approach. Todorov and colleagues (e.g., Todorov, 2011; Todorov, Pakrashi, & Oosterhof, 2009; Todorov, Said, Engell, & Oosterhof, 2008) concluded that human faces are evaluated on two dimensions: valence and power, two dimensions that are consistent with Fiske and colleagues’ dimensions of social group evaluations (Ames, Fiske, & Todorov, 2011). In laboratory explorations of these dimensions, Oosterhof and Todorov (2008) asked participants to generate trait descriptions of emotionally neutral faces. A principal component analysis of these descriptions revealed two components: trustworthiness (which accounted for 63% of the variance), which the authors labeled “valence,” and dominance (18% of the variance).

**Application of the Two Dimensions to Relationships**

Friendship and relationship researchers also identify ability/competence and trustworthiness/warmth as central to relationship-related judgments. Neuberg and colleagues (e.g., Cottrell, Neuberg, & Li, 2007; Neuberg, Smith, & Asher, 2000), for example, proposed a “sociofunctional analysis” that posited two dimensions critical for making judgments about others’ relationship potential: (a) trust and cooperation, which allows for interdependent relationships to develop, and (b) the relevance and quality of the other person’s characteristics, in which individuals differentially value another’s characteristics to the extent that they are relevant and
positive. Repeated tests of their sociofunctional analyses have supported such an approach, such that trust and cooperativeness were preferred across all types of relationships, but other characteristics (e.g., physical attractiveness, intelligence, leadership skills) were valued to the extent to which they were relevant to the type of relationship (e.g., Cottrell & Neuberg, 2005; Cottrell et al., 2007). A second approach, the Ideal Standards Model (Fletcher & Simpson, 2000; Fletcher, Simpson, & Thomas, 2000), describes three dimensions of relationship preferences that are generated from evolutionary theory and supported by empirical investigations: (a) intimacy and commitment, which assesses an individual’s ability to be a devoted parent and to form cooperative unions; (b) health and physical attractiveness, which assesses an individual’s overall health and fitness levels (and in the case of women, fertility); and (c) status and resources, which emphasizes an individual’s ability to ascend the social hierarchy and gain valued resources. The intimacy and commitment dimension reflects an assessment of a target’s trustworthiness/warmth, and the health and status dimensions reflect ability/competence assessments, the importance of which varies by sex and time. Third, Ellis (1992; 1998) developed a model that emphasizes the target person’s “ability and willingness” to fulfill gender-specific evolutionary-dependent needs. For instance, for women’s preferences in men, ability refers to a man’s earning capacity, and the willingness assessment refers to his desire to provide such resources to the woman. Likewise, multiple models for the matching effect (i.e., the tendency for individuals to pair up with others of similar levels of physical attractiveness) propose that two processes are operating: one that assesses the quality of the other (via the other’s physical attractiveness) and a second that assesses the degree to which the other would accept the individual (e.g., Huston, 1973; Huston & Levinger, 1978; Shanteau & Nagy, 1976, 1979).

Finally, friendship preferences also reflect dimensions consistent with a two-dimensional approach. Vigil (2007; 2009), for instance, posited that friendship choices are determined by (a) trustworthiness, which was defined as the other person’s willingness to engage in a mutually beneficial relationship and (b) capacity, which was defined as the other person’s ability to provide resources (e.g., financial resources, status) to another person. Repeated tests by Vigil (2007) revealed that, indeed, participants differentiated between capacity and trustworthiness cues, and that such differences predicted friendship preferences.

Summary

Though person perception and relationship researchers have used different labels to describe the two dimensions (e.g., competence, power, morality, warmth), across laboratory, research topic, and target group, research has consistently identified an evaluation of the other person’s (a) capability to facilitate or thwart one’s goals (e.g., competence, power) and (b) willingness to facilitate or thwart one’s goals (e.g., morality, warmth) as the components of interpersonal judgments. In addition, these dimensions dominate such judgments, with researchers concluding that these dimensions account for between 75% and 97% of the variance of overall evaluations of other people (Wojciszke, 1994; Wojciszke, Bazinska, & Jaworski, 1998).

A Two-Dimensional Approach to Interpersonal Attraction

The purpose of this article is to propose that the two-dimensional framework used by person perception researchers provides an approach to understanding the processes that underlie attraction. Consistent with models of person perception, we assume that perception serves an adaptive function in facilitating an individual’s interests and reproductive goals (Fiske, 1992; Kunda, 1999; McArthur & Baron, 1983; Wojciszke, 2005). From such a perspective, all animate life, from single-cell amoeba to human beings, have developed mechanisms to regulate movement away from possibly harmful environmental stimuli and toward possibly beneficial environmental stimuli (e.g., Kenrick & Shoda, 2008; Schneirla, 1959). In this way, the decision to approach or withdraw from an environmental stimulus is a fundamentally adaptive decision that all organisms make. Given this logic, the two-dimensional model of attraction posits that such approach/avoidance decisions (i.e., behavioral attraction), and the affective responses (affective attraction) that often accompany them, flow from the trustworthiness/warmth and ability/competence assessments identified by person perception researchers.

The two-dimensional model further posits that a perceiv-er’s goals and interests are of fundamental importance to his or her attraction to a specific person. Goals and interests vary in importance across time and situation, with some being relatively high in motivational priority and other goals being less so. However, some goals, such as social acceptance, self-interest, or avoidance of pain (Fiske, 1992; Wojciszke, 2005) are, by comparison to other goals, consistently high in motivational priority across time and situations. Other goals, such as achievement or reproductive goals (e.g., Maner, Miller, Moss, Leo, & Plant, 2012), are more time- and situation-specific. By this reasoning, when no situation-specific goal has gained motivational priority, one’s chronic goals maintain their relative motivational priority and guide assessments of other people along the dimensions of trustworthiness/warmth and ability/competence (Fiske & Neuberg, 1990), and attraction should result from these relatively general evaluations. However, when a situation-specific goal is activated or salient, individuals are motivated to make specific evaluations to determine the extent to which the other person facilitates or hinders the activated goal (Fitzsimons & Shah, 2009). Such chronic and situation-specific goals can affect attraction whether the goal is explicit (consciously
recognize the influence of the perceiver’s goals and needs (see Kelley, 1971, 1972). In this article for clarity and capacity for the ability/competence dimension. We selected these terms because they speak to assessments of another person’s willingness and capacity to act in a particular way relative to an activated or chronic goal. When it comes to an attraction assessment toward a specific person, person perception terms (e.g., warmth, morality) are not ideal because they describe global evaluations of people rather than how another person might specifically respond to the perceiver. That is, when it comes to attraction to a specific person, a target person’s level of “morality,” for example, is not of primary concern. Morality is relevant to attraction only to the extent that it leads the perceiver to see the target as potentially willing to facilitate the perceiver’s goals and needs (see Kelley, 1971, 1972). In the following section, we discuss evidence for the influence of assessments of willingness and capacity on attraction.

**Capacity**

The capacity assessment refers to the degree to which an individual is evaluated as possessing the qualities necessary to facilitate (or hinder) the perceiver’s goals and interests. The two-dimensional model proposes that people evaluate others with respect to the dimensions critical for facilitation of activated goals. Evidence that attraction depends on such goal-activated judgments comes from assessments of attraction in the face of shifting goals. As noted earlier, whereas some goals are relatively consistent and stable within and across people (e.g., need for acceptance), people’s goals vary considerably. That is, goals often change over time and across situations, which changes the qualities of other people that are relevant to those goals and, thus, affect attraction. Empirical efforts to assess attraction in the face of changing goals provide compelling evidence for the importance of capacity assessments to interpersonal attraction. Such efforts have investigated attraction responses (a) across different types of relationships, (b) relative to shifting motivations of activated goals, (c) within individuals whose goals within a particular relationship type change across time, and (d) between individuals whose common goal is served best by different characteristics.

First, people have different goals in different relationships, and such differing goals are met most effectively by other people with different characteristics, suggesting that attraction should change accordingly. For example, when considering a potential friend, undergraduate students prefer those who offer intimacy/warmth rather than who excel academically. When considering who they want in a leader, however, preferences are reversed (Lusk, MacDonald, & Newman, 1998; MacDonald, 1998). Alternatively, when considering potential mates, kindness is more important than leadership potential, but not when thinking about a coalition partner (Dekay & Buss, 1996; cited in Buss, 1996). In this way, attraction to intimate/kind individuals as friends or mates reflects their capacity to meet the relationship-specific needs of intimacy and social/emotional support, and the preference for an accomplished leader or coalition partner flows from the relationship-specific goal of group task performance. Similarly, individuals prefer emotional closeness more in romantic partners compared with same or other sex friends (Fuhrman, Flanagan, & Matamoros, 2009) and “feel better” sharing material resources with siblings than strangers (Ackerman, Kenrick, & Schaller, 2007; see also Lieberman, Tooby, & Cosmides, 2007). In a similar vein, whereas trustworthiness (a trait synonymous with a willingness-to-benefit assessment) was highly desired across different types of relationships, preferences for capacity-related traits differed as a function of the demands of the specific relationship (Cottrell et al., 2007). In sum, as relationships change, so do one’s goals, and the evaluations of others are based more strongly on their capacity to facilitate those goals (also see Abele & Brack, 2013).

Second, research on motivated cognition indicates that evaluations of objects and people are affected by the degree to which they are viewed as aiding or inhibiting one’s goals (Balcetis & Dunning, 2006; Bargh, Gollwitzer, & Oettingen, 2010; McCulloch, Ferguson, Kawada, & Bargh, 2008). When a goal is activated, persons and objects that were viewed as able to facilitate one’s goal were evaluated more favorably than those that were not (Ferguson & Bargh, 2004). Furthermore, significant others who had the capacity to facilitate one’s interests, compared with those who could not, were evaluated more favorably and brought to mind more quickly (Fitzsimons & Shah, 2008). In addition, research has also revealed that people classify individuals into the social categories of “useful” versus “not useful” relative to their activated goal (Fitzsimons & Shah, 2009).

Third, which goal or interest is activated not only differs across relationship or individual, but also differs within the same person across time. As an example, Gangestad and Thornhill (1997) investigated women’s attraction to other people as a function of different days of their ovulatory
cycle. The researchers reasoned that evolutionary goals regarding genetic fitness are particularly pronounced when women are ovulating (e.g., when reproduction is most likely). Alternatively, near the beginning or end of the cycle, such interests should be less pronounced. Consistent with this reasoning, women who were ovulating expressed a stronger preference for socially dominant men than did women who were not ovulating, but only when considering the men’s attractiveness for a short-term relationship (i.e., one-night stand). Women also exhibit a stronger preference for symmetrical faces (e.g., Thornhill & Gangestad, 1999) and masculinized faces (e.g., Johnston, Hagel, Franklin, Fink, & Grammer, 2001) when they are high in conception risk. That is, differences in attraction flow from the manner in which reproductive goals wax and wane throughout the menstrual cycle. When reproductive goals are active, the quality of the man’s genes is more critical, and women express a greater preference for cues to genetic fitness.

Finally, the link between the capacity assessment and attraction is evidenced in the attraction patterns of groups of individuals who share a goal but for whom the goal can be satisfied via different strategies. For example, mate preferences are driven by a common goal that is pursued differently by men and women. According to theories of natural and sexual selection (Darwin, 1859; Symons, 1979), preferences for specific attributes in a potential mate reflect the fundamental drive to increase genetic fitness and reproductive success. This common goal is served best by different strategies in men and women. For instance, physical cues that suggest a woman’s fertility should be paramount in men’s mate preferences, and indeed men, more than women, exhibit preferences for an optimal waist-to-hip ratio (D. Singh, 1993), cues to youth (Buss, 1989a; Thiessen & Gregg, 1980), and physical attractiveness (Buss, 1989a). Alternatively, women’s mate preferences focus more on a man’s ability to care for her and her offspring. Women, compared with men, prefer strength, status, and potential to acquire resources (Buss & Barnes, 1986; Buss & Schmitt, 1993). In short, people differ in the manner in which a goal is fulfilled and are more attracted to those who are evaluated as capable of meeting that goal.

**Summary.** Overall, the empirical evidence is clear that attraction tendencies change with the assessment of a target’s capacity to meet activated goals. Such goals change across relationships, individuals, and time and in turn, so does attraction to different targets. Such evidence provides support for the sensitivity of attraction to active goals and the crucial role that the capacity assessment plays in determining attraction to another person.

**Willingness**

The willingness assessment refers to the degree to which an individual is evaluated as willing to potentially facilitate the interests/goals of the perceiver. Put another way, this assessment reflects one’s belief that a target will act benevolently toward the perceiver in the future. This definition of willingness aligns closely with definitions of trust: Frost, Stimpson, and Maughan (1978) described trust as the expectancy that another will act benevolently toward the person. Similarly, Kelley and Thibaut (1978) defined trust as the “assurance of not being exploited” during a social interaction (p. 232), and Pruitt and Smith (1981) defined trust as confidence that the other person is “cooperatively motivated.”

Previous definitions and measures of warmth/morality are consistent with this conceptualization of willingness: Fiske and colleagues’ (Fiske, Cuddy, & Glick, 2007) “warmth” is defined as the perceived intent of another group as able to help or harm the in-group, and is operationalized using traits such as “moral,” “trustworthy,” “sincere,” and “helpful”; Peeters and colleagues’ “other-profitability” defines the target’s perceived adaptive value to others (Peeters & Czapinski, 1990), and is operationalized using traits such as “tolerant,” “trustworthy,” and “trust” (as used in Wojciszke). Peeters and colleagues’ (as summarized in Wojciszke, 2005) trust is defined as the ‘moral’ of a person’s goals relative to moral norms, and operationalized using “kind,” “noble,” “true,” and “right.” Given this foundation in the person perception literature, the assessment of willingness is best viewed as an assessment of the extent to which a target is viewed as willing to facilitate one’s interests.

Despite the considerable consensus regarding the definition of trust, research has not always differentiated between trust and its “near neighbor” constructs, such as predictability, confidence, loyalty, cooperativeness, and honesty. These related concepts all concern whether another person will act in a trustworthy fashion during a future interaction (Mayer, Davis, & Schoorman, 1995). Trust, however, appears to be a more general term than these other constructs. Indeed, given different contexts, trust takes on a more specific meaning (e.g., loyalty, confidence). For example, in a boss–employer relationship, general trustworthiness assessments are indicative of another’s “dependability,” but in a fraternity, general trustworthiness assessments are characterized as another’s “loyalty” (see Cottrell et al., 2007). Unfortunately, trust is often confused with—or is not empirically differentiated from—these related concepts, including predictability (Dasgupta, 1988; D. Good, 1988), cooperation (Axelrod, 1984), contractual obligation (Young & Wilkinson, 1989), confidence (Coleman, 1990; A. P. Jones, James, & Bruni, 1975), and even reliability/credibility (for a discussion, see Blomqvist, 1997). However, these related terms and concepts are perhaps best viewed as context-specific manifestations of a general trust construct.

By this reasoning, the willingness-to-benefit assessment is a function of specific judgments on trait dimensions such as warmth and trustworthiness, all of which bear on the extent to which a target will be perceived as willing to invest time and energy to facilitate one’s interests. Importantly, whereas a capacity assessment depends more on the context (e.g.,
basketball skills versus reading ability when selecting a lab partner), the willingness assessment reflects a relatively general and stable belief that the other person is potentially inclined toward facilitating interests even outside of the context in which that judgment is made (Fiske et al., 2007; Peeters, 2002).

**Willingness and attraction.** There is widespread agreement regarding the critical role that a willingness-to-benefit assessment plays in the development and maintenance of harmonious social relationships (e.g., Brewer, 1997, 1999; Holmes & Rempel, 1989; Simpson, 2007). People must engage in exchanges with unacquainted others, with others for the purpose of mating, as part of hierarchical relationships, and as part of intragroup and intergroup interactions. In each context, belief in other person’s potential willingness to facilitate beneficial outcomes is critical for the success of the social relationship, and thus, is important to attraction to another. More specifically, Cottrell et al. (2007) proposed that trust is desired in relationship partners because it predicts who will help with the acquisition of valued resources. In a similar vein, an evolutionary approach posits that we are attracted to those we perceive as cooperatively motivated because it signals their willingness to form a social relation (Gangestad & Simpson, 2000; Hazan & Diamond, 2000).

Empirical evidence supports the importance of willingness-to-benefit for facilitating successful exchanges across a broad spectrum of interactions, including ad hoc exchanges (Cosmides & Tooby, 1989), close relationships (e.g., Rempel, Holmes, & Zanna, 1985; Rempel, Ross, & Holmes, 2001), and intergroup relations (Wildschut, Pinter, Vevea, Insko, & Schopler, 2003). In fact, in a cross-cultural sample, people from different countries consistently displayed greatest preference for values related to another’s benevolence (e.g., honesty, loyalty, helpfulness, forgiveness; Schwartz, 1992; see also, Schwartz & Bardi, 2001). In the mate selection literature, research has identified a clear preference for trust-related traits (e.g., Buss & Barnes, 1986), and across a variety of different relationship types, trust-related traits are highly desired (Cottrell et al., 2007). In short, theory and evidence point to the importance of willingness-to-benefit assessments to attraction across relationship types and cultures.

**Determining Attraction From Capacity and Willingness Assessments**

We posit that specific cues to another’s capacity and willingness-to-benefit are integrated to generate an overall capacity and an overall willingness assessment. Consistent with the theorizing of Brunswik (1956), Buss and Schmitt (1993), and others (Miller, 2000; Symons, 1995), cues to another’s capacity and willingness are weighted subjectively relative to activated goals. Attraction to a target results from the integration of these individual assessments.

For instance, when a prospective dating partner possesses desired attributes (e.g., physical attractiveness, sociability, and creativity), one might evaluate the partner as highly “capable” of meeting one’s goals. At a more micro level, people are able to arrive at a single physical attractiveness assessment despite the vast phenotypic variability that occurs between people (e.g., differences in facial attractiveness, symmetry, body scent; see Grammer, Fink, Juette, Ronzal, & Thornhill, 2002). The positive consequences of the physical attractiveness assessment (e.g., genetic fitness, social status) would be weighed against the negative consequences of physical attractiveness (e.g., stereotypes regarding lower ability, “beauty is beastly”; Heilman & Stopeck, 1985), and would be combined with the weighted contribution of other capacity assessments (e.g., sociability and creativity). The weights applied to each characteristic would differ as a function of the perceiver’s goals and interests and cultural influences (e.g., F. Marlowe & Wetsman, 2001; Symons, 1995) and would produce an overall capacity assessment. The same process occurs with the willingness assessment; however, this assessment is less tied to situational constraints than is the capacity assessment. The positivity of the capacity and willingness-to-benefit assessments determines the nature and amount of affective and behavioral attraction. As a general rule, the more favorable the assessments, the more attraction one experiences.

This process does not necessitate explicit, consciously processed, or direct information about a target related to capacity and/or willingness. Individuals can generate capacity and willingness assessments rapidly and automatically from the available information (e.g., Ferguson & Bargh, 2004; Oosterhof & Todorov, 2008). In addition, individuals infer “missing” willingness information from nothing more than a simple photograph of another (Condon & Crano, 1988) and infer less willingness from an attractive, versus an unattractive, potential partner after only a brief encounter (S. Kiesler & Baral, 1970). In this way, perceivers can experience affective and behavioral attraction to individuals about whom there is little explicit, conscious, or direct information regarding the other person’s willingness or capacity.

**Relation between willingness and capacity assessments.** The two-dimensional model states that an attraction assessment can be predicted by a weighted combination of the willingness-to-benefit and capacity assessment. Stated algebraically, attraction can be estimated by

\[
IA_p = \frac{W}{|W| + |C|} - \frac{W_{ij}}{w_i + w_j}
\]

such that \(IA_p\) is interpersonal attraction toward person \(p\) (in which positive values indicate attraction), \(W\) is the
The second term on the right side determines the strength of the response. First, the weight for willingness-to-benefit and capacity is determined by (a) the positive or negative valence of the descriptor, such that greater weight is given to negative attributes (R. Singh & Ho, 2000; R. Singh, Lin, Tan, & Ho, 2008; R. Singh & Teoh, 1999), and (b) for the capacity assessment, the relevance of each individual descriptor to the demands of the interaction (e.g., basketball skills and physical attractiveness should be weighted differently when considering a potential basketball teammate or dating partner; Cottrell et al., 2007; Lusk et al., 1998; MacDonald, 1998). Second, the willingness assessment's weight \( w_1 \) is greater than the weight for capacity assessment \( w_2 \). From an evolutionary perspective, an assessment of the other person's intentions to help or harm is more important to one's reproductive potential as compared with the other person's ability to act on this intention (Fiske et al., 2007; Cacioppo et al., 1997; Peeters, 2002).

**Conclusion**

We have proposed that the two dimensions by which others are evaluated also provide the basis for interpersonal attraction. The evidence for this model is compelling, as it includes work from a variety of different laboratories, on a diversity of topics, and that originated without the intent of speaking to the current model. The importance of capacity assessments for attraction is supported effectively by investigations on mate preferences, menstrual cycle changes in attraction patterns, and leadership preferences, all of which assessed conditions in which changes in an active goal led to changes in attraction patterns. For its part, the importance of the willingness-to-benefit assessment is evidenced by the ever-growing literature that establishes trust as a critical predictor of human relationships and, as such, a factor to which people are particularly sensitive. Our next goal is to assess the merit of this model by evaluating it relative to the extant literature on attraction.

**Determinants of Attraction**

An important test of the merit of the two-dimensional model is the extent to which it can account for reliable findings within the attraction literature. Over the past 50 years, reviews of the interpersonal attraction literature (e.g., Berscheid & Walster, 1969; Finkel & Baumeister, 2010; Huston, 1974; Orbuch & Sprecher, 2003; Simpson & Harris, 1994) have consistently focused on the most robust and extensively researched attraction phenomena, including the reciprocity effect, pratfall effect, matching hypothesis, influence of arousal, and similarity effect. In the following section, we discuss how each of these phenomena can be understood in terms of the two-dimensional model.

**Reciprocity Effect**

Extensive scholarship, beginning with Gouldner’s (1960) seminal treatise, has demonstrated that we like those who express liking for us. Reciprocated attraction (hereafter referred to as the reciprocity effect) has been demonstrated between groups (Burleson, 1983) and individuals (e.g., Wilson & Henzlik, 1986). A meta-analysis, which included 17 studies manipulating the presence of ingratiation on attraction, found a moderate and positive effect of expressed liking on reciprocated attraction (Gordon, 1996; see also Kenny, 1994).

The initial explanation for the reciprocity effect was posited by Gouldner (1960), who argued for a norm of reciprocity mechanism. In its simplest form, the norm of reciprocity posits that individuals should help, or at least do no harm to, those who have helped them. Gouldner generalized the receipt of aid to receiving a benefit in the form of another’s affections. If someone expresses affection for another, that affection activates the reciprocity norm, and the recipient reciprocates the affection.
From the perspective of the two-dimensional model, reciprocated attraction operates via changes in willingness-to-benefit information. Montoya and Insko (2008), for example, proposed that learning that one is liked by another person leads to attraction—not because it activates a reciprocity norm—but rather, because the received attraction conveys the admirer’s intent to act in a trustworthy and benevolent fashion during a future interaction. This point was also made by Ellis (1998), who proposed that symbolic acts of investment (e.g., “I love you.”) convey cues to the individual’s willingness to continue investing in the relationship. Similarly, the commitment model (Frank, 1988, 2001) argued that cooperative systems between individuals are facilitated by having identifiable marks—including nonverbal behaviors such as blushing or expressions of positive emotions—to help differentiate those who would cooperate from those who were less likely to cooperate.

This interpretation of the reciprocity effect suggests two empirical effects: (a) expressed attraction leads to expectations of benevolence in the admirer (i.e., expressed attraction leads to a more positive willingness-to-benefit assessment) and (b) expectations of benevolence lead to reciprocated attraction to the admirer (i.e., favorable willingness-to-benefit assessments predict attraction). With regard to the first suggestion, research indicates a relation between the receipt of attraction from a person and perception of that other person’s trustworthiness (Hawes, Mast, & Swan, 1989; Swan, Trawick, & Silva, 1985). Specifically, participants infer more trustworthy intentions from those who express liking toward them. Evidence is also consistent with the second suggestion: Individuals like those they view as benevolently oriented. Doney and Cannon (1997) concluded that perceptions of salespeople’s benevolence during buyer–seller interactions were critical to determining the degree to which salespeople were liked. Experimental investigations of the impact of benevolence also confirm that individuals experience more liking for those from whom they expected future benevolence compared with those without such an expectation (Ayers, Nacci, & Tedeschi, 1973).

A direct test of the mediational role of willingness-to-benefit in the reciprocity effect also supports the two-dimensional model. Montoya and Insko (2008) manipulated the amount of expressed liking a participant received from a target and then measured attraction to and perceived benevolent intent in the target. The researchers found that the expression of liking led to more attraction, and that the perception of benevolent intent mediated the attraction generated by the expression of liking. In total, the willingness-to-benefit assessment successfully accounts for the reciprocity effect: Expressed attraction leads to perceptions of willingness to facilitate future interests and goals, and willingness facilitates reciprocated attraction.

Importantly, if the reciprocity effect is a function of a change in the willingness assessment, we should expect factors that inhibit the reciprocity effect to do so by reducing willingness-to-benefit assessments, and indeed, research has supported this expectation. E. E. Jones (1964), in his book on ingratiation, discussed numerous limitations and facilitators of the reciprocity effect. The list included many of the most pervasive real-world concerns associated with the receipt of an expression of liking, including the amount of sacrifice the other person has made for the admired, the perceived sincerity of the expressed liking, the presence of ulterior motives for the expressed liking, among others. In the section that follows, we explore each concern and its relevance to the reciprocity effect in light of a willingness evaluation.

**Amount of sacrifice.** The reciprocity effect is more potent if an individual perceives that a target person has made a costly sacrifice of his or her self-interest (e.g., Nadler, Fisher, & Streufert, 1974). For example, in an experimental investigation of the importance of sacrifice, participants interacted with a confederate who had resources of either $1 or $4 (vs. the participants’ $2) and gave participants either 20% or 80% of their money. Participants who received 80% of the confederate’s $1 gave more money back to the confederate and were more attracted to the confederate than were participants who received 20% of $4 (Pruitt, 1968). Although all participants received the same amount of money, attraction was greater when the participants perceived the gift as costly.

From the perspective of a two-dimensional model, people are more likely to perceive a person who gives up a great deal to (or as a part of) an exchange with them as willing to facilitate beneficial interactions in the future and, thus, as more attractive. This notion is also consistent with an evolutionary biological perspective, which suggests that perceived sacrifice operates as a cue to the other’s honesty. Such sacrifices are evaluated to be honest if, and only if, they are costly to display (Zahavi, 1975, 1977). If a display is presented without cost, the display is unreliable because it is too easily mimicked and the chances of being cheated increase (see also Grafen, 1990; Zahavi & Zahavi, 1997). Thus, it is not just sacrifice that breeds attraction; only sacrifice that points to a potential willingness to facilitate one’s interests breeds attraction.

Other studies have similarly supported the importance of the willingness-to-benefit assessment in the face of sacrifice. For instance, experimental work has found that participants experienced the greatest amount of attraction for confederates who offered assistance that implied, rather than did not imply, rewarding future interactions and positive outcomes (Nadler et al., 1974). Similarly, attraction to those who offered costly, rather than non-costly, aid covaried with the expected willingness of their partner to provide positive interactions in the future (Fisher & Nadler, 1976). Further support for the two-dimensional model comes from researchers who have theorized that attraction is reciprocated to the extent that the other’s actions were viewed as predictive of another’s willingness to facilitate desired future outcomes (e.g., E. E. Jones,
Jones, & Gergen, 1963; Nemeth, 1970; also see Schopler, 1970). In each of these situations, attraction increased in so far as sacrifice and/or assistance indicated that the partner was willing to facilitate the participant’s future interests.

**Perceived sincerity and ulterior motives.** The power of the reciprocity effect is also affected by the perceived motives of the admirer; in so far as motives are viewed as sincere, attraction is reciprocated. In the absence of perceived sincerity, attraction is less likely to be reciprocated. The dominant interpretation of this moderating effect is psychological reactance (Brehm, 1966). In Brehm and Cole’s (1966) classic study, participants in a high-importance condition were told that the study’s goal was to predict the other participant’s future success in life and that the accuracy of one’s prediction was critical. Participants in the low-importance condition were simply told the study’s goal was the completion of a student project. Next, the other participant (actually a confederate) either brought the participant a soda (favor) or not (no favor). Results revealed that participants in the high-importance condition were less likely to help the confederate after receiving a favor. According to the reactance explanation, the favor in the high-importance condition threatened participants’ freedom to evaluate the confederate accurately; freedom that they worked to reestablish by reducing evaluations of the confederate.

However, Worchel, Andreoli, and Archer (1976) offered a different explanation, one that is consistent with a focus on the importance of a willingness assessment. Their study replicated Brehm and Cole’s (1966) procedure but also manipulated the attributions participants could make after the high-importance versus low-importance manipulations. The researchers noted that perceptions of ulterior motives were greatest in the high-importance condition and covaried with attraction assessments. In other words, the admirer was liked to the extent that ulterior motives for the expression of liking could not be made. From the perspective of the two-dimensional model of attraction, the perception of ulterior motives for the gift undercuts the assessment of the admirer’s willingness to facilitate future goals/interests and thus, diminishes attraction to the admirer. After all, a gift given in hopes of receiving a situation-specific outcome is less likely to indicate a situation-general disposition toward benevolence, and it is the perception of just such a disposition that leads to attraction. Other researchers have demonstrated empirically that assessments of the expected positivity of future interactions with a partner who did or did not have ulterior motives covaried with reciprocated attraction (e.g., Brounstein, Norman, & Ostrove, 1980; E. E. Jones, Stires, Shaver, & Harris, 1968; Nadler, Fisher, & Ben Itzhak, 1983; Schopler & Thompson, 1968). And additional empirical work has confirmed these findings by noting the importance of perceptions of sincerity to reciprocated attraction (see Goranson & Berkowitz, 1966; Greenberg & Frisch, 1972; Greenberg & Saxe, 1975).

**Summary.** The two-dimensional model posits that an expression of liking provides information about the other person’s willingness to facilitate the person’s interests. Such an assessment of willingness leads to reciprocated attraction. Any information that is indicative of liking (e.g., a warm smile, giving flowers, saying “I love you.”) predicts attraction via an augmented willingness assessment. In the absence of such a perception of willingness—as in when the sincerity of the expression of liking is in doubt due to ulterior motives or lack of sacrifice—reciprocated attraction is attenuated or absent. The empirical evidence is consistent with this expectation: Research on the reciprocity effect notes (a) a link between perceived liking and an inference of willingness to meet future goals/interests, (b) a link between that inference and reciprocated attraction, and (c) moderators whose potency is tied to their ability to undercut that inference (for a more complete discussion, see Montoya & Horton, 2012). Contrary evidence for the two-dimensional model would be found if (a) willingness were to fail to mediate the acceptance-to-reciprocated attraction link, (b) expressions of attraction that do not inform perceptions of the other’s actual beliefs (e.g., flattery deemed insincere) were to result in reciprocated attraction (for a discussion, see Gordon, 1996), or (c) positive affect or the capacity assessment were found to mediate the reciprocity effect above the effect of the willingness assessment. Specific to this last point, research has failed to find evidence that the reciprocity effect is mediated by self-esteem or vanity (Vonk, 2002; Studies 1-4), positive mood (Study 3), or narcissism (Study 4).

**Desirable Characteristics**

Considerable evidence across research domains indicates that people prefer to affiliate with others who possess particular attributes, including health (Ford & Beach, 1951; Thiessen & Gregg, 1980), intelligence (Buss, 1989a), and good earning capacity (Buss & Barnes, 1986). The two-dimensional model posits that a characteristic leads to attraction when it is perceived to make a target more capable of facilitating the perceiver’s interests. In other words, positive characteristics (e.g., physical attractiveness, intelligence) are considered to convey information about capacity. However, a critical test of the two-dimensional approach is to explain situations in which such characteristics do not lead to attraction. Fortunately, research has explored situations in which attraction is not greater for those who possess certain highly regarded attributes, and in this context, we evaluate the evidence relevant to the two-dimensional model.

**Pratfall effect.** As just noted, favorable capacity assessments (in the form of “desirable characteristics”) are often linked to increased attraction, but this is not always the case: For example, group followers do not always select the most capable individual as their leader (Radloff & Helmreich, 1968), the most competent individuals are not always rated
the most likable (Holland & Webb, 1955), and sometimes the "lovable fool" is preferred over the "competent jerk" (R. Singh & Tor, 2008). A simplistic view of a two-dimensional approach might regard such findings as inconsistent with predictions in that increasing one’s capacity assessment (e.g., "The person is a ‘high-quality person.’") should result in more attraction. However, such findings are consistent with the two-dimensional model when one considers the impact of the target’s competence/quality on the willingness assessment. Specifically, when interacting with an exceptional other, concerns regarding whether the other will accept the perceiver (i.e., be high in willingness-to-benefit) will arise. In such cases, the perceiver’s attraction should decrease even while capacity evaluations (i.e., how capable, talented the other was) remain favorable.

Research on the pratfall effect is consistent with this notion. Aronson et al. (1966), in their famous "pratfall study," found that participants were more attracted to a competent target when that target committed an embarrassing blunder than when the target did not commit the blunder. Aronson and others explained the finding by suggesting that a blunder made the competent person "seem less austere, more human . . . " (p. 228); in other words, the blunder decreased the perceived probability that the perceiver would be evaluated negatively by the target (Tesser, 2000; see also C. Kiesler, Weizmann, & Palla, 1967). Helmreich, Aronson, and LeFan (1970) later reasoned that the blunder "humanized" the confederate and made him more approachable. In this way, a blunder’s impact on attraction resulted from a change in the perceived probability that the target would positively evaluate the participant.

More recent research has observed a similar effect of the target person’s qualities on attraction. An exceptionally capable target may be assessed as low in willingness-to-benefit and thus, the perceiver would experience little affective and behavioral attraction toward the target while continuing to assess the target as highly capable. Herbst, Gaertner, and Insko (2003), for example, investigated affective responses to a potential partner whose ability level on a dimension varied in relation to the participants’ actual and ideal self. When participants expected to meet a partner who exceeded their ideal self, they evaluate their partner’s ability favorably, but their affective responses were less favorable than they were to less-exceptional partners or to exceptional partners who they thought they would never meet. Similarly, Montoya and Horton (2004) manipulated whether or not participants thought they would meet a partner and the perceived ability level of the partner. The evaluation of their partner’s attributes (i.e., the capacity assessment) became more positive as ability increased. Participants expressed less behavioral attraction toward "exceptional" partners, but only if they expected to meet them face to face. Amabile (1983) observed a similar pattern, such that although participants regarded individuals who criticized others as more intelligent, these individuals were also rated as less likable.

In short, the willingness assessment often becomes less favorable in the face of exceptional others due to the perception of the greater probability of rejection (while capacity assessments remains favorable), and reductions in attraction follow suit.

In this way, individuals who are blessed with highly favorable characteristics (e.g., high intelligence or exceptional beauty) are perhaps not always the target of tremendous behavioral attraction (at least relative to the affective attraction they inspire), as the positivity of their attributes leads observers to infer less willingness-to-benefit.

**Matching hypothesis.** Despite the overwhelming evidence that people are most attracted to those who are highly physically attractive, people end up actually pairing up with others who more-or-less match their level of physical attractiveness. This tendency, referred to as the matching effect (Walster, Aronson, Abrahams, & Rottman, 1966), is a robust phenomenon, having been documented numerous times (Cash, 1981) and having held up under meta-analytic scrutiny. Indeed, a meta-analysis of 18 matching studies found a reliable correlation between the physical attractiveness of relationship partners (r = .39; Feingold, 1988). According to the original explanation for this effect (Walster et al., 1966), a perceiver’s self-evaluation is critical for differentiating between realistic mate choices—choices determined by the objective desirability of the date while constrained by the perceived likelihood of attaining the date—from idealistic mate choices—choices that are determined solely by the objective desirability of the date. “Realistic” individuals consider the potential partner’s physical attractiveness and the probability of acceptance, and then attempt to date the most attractive other who will accept them. Thus, realistic individuals, due primarily to equity and market processes, end up paired with others who approximate their own physical attractiveness.

Embedded within the above reasoning is the notion, as with the pratfall effect, that people assume that highly attractive others (i.e., highly capable others) are more likely to reject them (i.e., will be low in willingness-to-benefit) than are less-attractive others. Indeed, within the romantic partner domain, assessments of capacity and willingness are inversely related: Participants regard attractive targets, either directly or by inference, as more capable of meeting their needs but less willing to do so (see Dion, Berscheid, & Walster, 1972).

In this way, the matching effect is consistent with the two-dimensional model’s description of the role of willingness-to-benefit assessments and the impact of perceived rejection on that assessment. Indeed, a number of theorists (e.g., Huston & Levinger, 1978; Shanteau & Nagy, 1979) have made a similar argument that matching is essentially comprised of two independent assessments: an assessment of attractiveness (capacity assessment) and an assessment of the likelihood of acceptance (willingness assessment).
Models of risk management systems (Murray, Holmes, & Collins, 2006; Murray, Holmes, & Griffin, 2000) have echoed a similar sentiment, positing that individuals in relationships consider risks of rejection alongside relationship promotion goals. Risk management models argue that the belief and confidence in another person’s liking indicates that it is safe to invest further in the relationship.

However, the two-dimensional model goes further to predict a divergence between people’s willingness-to-benefit and capacity assessments and their attraction to highly physically attractive targets. Such targets, who would be regarded as highly capable of meeting one’s goals but less willing to do so, should produce relatively high affective attraction but relatively low behavioral attraction. Although a distinction between willingness/capacity and attraction has been documented numerous times in the pratfall effect literature, to our knowledge only one study has identified this differentiation specifically for the matching hypothesis: Montoya (2008) found that participants’ assessments of a target’s physical attractiveness increased as the target’s objective physical attractiveness increased, but that expectations of rejection (willingness assessment) also increased and judgments of the likelihood of a relationship developing decreased.

Role of self-esteem. The two-dimensional model can also shed light on the important role of self-esteem in romantic behavior. From the perspective of the two-dimensional model, an individual’s self-esteem is associated with the way willingness-to-benefit information is interpreted. For example, Cameron, Stinson, Gaetz, and Balchen (2010) found that people felt more accepted by a confederate who smiled and laughed, maintained eye contact, and used animated facial expressions than one who smiled less, maintained little eye contact, and used few vocal inflections. However, when there was the possibility of rejection (i.e., an intense interaction with an attractive confederate), participants with low self-esteem reported fewer acceptance cues from the confederate, despite the presence of these same positive signs of willingness (i.e., smiling, eye contact, etc.). Such findings indicate that for individuals with low self-esteem, such behaviors are either missed or are not interpreted as signs of acceptance, which is associated with a less-favorable willingness-to-benefit assessment and thus, less attraction. In this way, individuals with high self-esteem would be likely to regard higher quality partners as more willing to meet their goals than individuals with low self-esteem.

Consistent with this notion, self-evaluations, which have been central to a diverse set of mate-preference models (e.g., Feingold, 1988; Huston, 1973; S. Kiesler & Baral, 1970; Sloman & Sloman, 1988; Walster et al., 1966), are hypothesized to be important because they represent one’s belief that one can achieve a desired goal, and in the case of attraction, whether one would be accepted and liked by the other. Such an interpretation is also consistent with an evolutionary perspective, which posits that in the dating context, one’s self-esteem reflects one’s self-perceived quality as a romantic partner (e.g., Dawkins, 1982; Kenrick, Groth, Trost, & Sadalla, 1993; Todd & Miller, 1999; Wright, 1994). Kirkpatrick and Ellis (2004) similarly proposed that one function of self-esteem is to facilitate approach toward others who are high in mate quality, but not so high that they would not reciprocate benefits (see also Kavanagh, Robins, & Ellis, 2010).

Empirical evidence from the matching literature is consistent with such theorizing. In the dating realm, individuals with an unfavorable self-evaluation believe it less likely that they can attract (i.e., be accepted by) a desirable target than do individuals with a favorable self-evaluation (Lewin, Dembo, Festinger, & Sears, 1944). Furthermore, men whose self-esteem had been experimentally lowered displayed more romantic approach behaviors toward a moderately physically attractive woman than toward a very attractive one. In contrast, men whose self-esteem had been experimentally raised show the reverse pattern (S. Kiesler & Baral, 1970). Similarly, participants with higher self-rated mate value express preferences for higher quality mates than do those with lower self-rated mate value (e.g., Buss & Emery, 2003; Pawlowski & Dunbar, 1999; Regan, 1998). Overall, the relation between self-esteem and attraction in the dating realm is consistent with the two-dimensional model’s notion that self-esteem is related to assessments of a target’s willingness to meet one’s relationship needs.

Summary. The two-dimensional model hypothesizes that individuals are attracted to those who possess characteristics that indicate a capacity and/or willingness to facilitate one’s goals and interests. Characteristics (and the individuals who are perceived to have them) are desirable because they reflect a capability to meet a perceiver’s goals and interests, and when there is no possibility for a negative evaluation (e.g., as part of a no-interaction task or a person-perception task), attraction responses should align with the capacity assessment. However, certain characteristics, such as intelligence or physical attractiveness, can have opposite effects on the capacity and willingness-to-benefit assessments, especially when the situational context includes the possibility of rejection or upward social comparison. As levels of such characteristics rise, the capacity assessment is likely to continue to rise, but the willingness assessment is likely to drop, taking with it attraction responses.

Indeed, investigations of the pratfall effect, the matching phenomenon, and the role of self-esteem in romantic partner selection provide consistent evidence that even seemingly universally positive traits, such as physical attractiveness and intelligence, are not always related positively and linearly with affective and behavioral attraction. The two-dimensional model, with its joint focus on the willingness and capacity assessments (and the interaction between the two), demonstrates how these forces work to produce attraction.
Presence of Arousal/Dissonance

In theoretically related but distinct lines of research, considerable work has found that dissonance and physiological arousal can be associated with more attraction. Research on these processes is divided into distinct lines of research, each with different methods, conclusions, and theoretical explanations. For purposes of clarity, we maintain those distinctions. However, we acknowledge that there is conceptual overlap between arousal and dissonance, in that arousal is a component of dissonance and dissonance can be misattributed to external sources (Fazio, Zanna, & Cooper, 1977). We use the two-dimensional model to explain how attraction can result from (a) the anticipation of negative-state (i.e., expecting painful electrical shocks in the near future) and (b) past negative-state experiences (i.e., a, a harsh initiation).

Anticipated future distress research. In the 1950s, researchers made the initial observation that anticipated exposure to a noxious stimulus increased participants’ behavioral attraction to another person who was to share the same noxious stimulus (Schachter, 1959). More succinctly, “Misery does not just love any kind of company, it loves only miserable company” (Schachter, 1959, p. 24). In an initial study that became the template for many studies that would follow, Schachter (1959) told some female participants that they would soon participate in a task that would involve receiving intensely painful electrical shocks (high-anxiety condition). He told other participants that they would receive shocks that would feel like “a tickle or a tingle” (p. 14; low-anxiety condition). Next, participants were given the option to wait either alone or with others who were to receive similar shocks. Participants in the high-anxiety condition, compared with participants in the low-anxiety condition, more frequently chose to wait with other participants who were to face similar shocks. Schachter identified five explanations for the affiliative desires of those in the high-anxiety condition, most notably, cognitive clarity, which referred to the desire to affiliate to understand the painful consequences that await them, and direct anxiety reduction, which referred to affiliating to secure comfort and assurance from others.

The two-dimensional model regards the Schachter paradigm as another example of how situation-specific shifts in goals lead to changes in attraction. That is, the Schachter paradigm changes the participant’s goals and thus, the qualities of others that are evaluated as most capable of meeting those goals. In the case of the threat of shocks, for example, the threat creates a desire for information and/or for comfort, resources that others who face a similar plight are most capable of providing. A review of the literature revealed 25 studies that examined the influence of anxiety on affiliation and confirmed the basic notion that anxiety breeds affiliative tendencies (e.g., Darley, 1966; Dembroski & MacDougall, 1978; Fox, 1980, 1981; Strümpfer, 1970; Zucker, Manosevitz, & Lanyon, 1968). Other research investigated situational factors that moderate participants’ experiences of attraction to others who might share a common fate. The typical procedure replicated Schachter’s basic paradigm and then either manipulated the participant’s goals or assessed attraction to different targets who were equipped to different degrees to meet an activated goal. For such studies, the two-dimensional model posits that anticipation of distress generates a desire, and that participants experience attraction to those who are seen as capable of satisfying that desire.

First, in a series of intriguing field studies investigating affiliative patterns of pre- and post-operative hospital patients, Kulik and colleagues (Kulik & Mahler, 1989; Kulik, Mahler, & Moore, 1996; Kulik, Moore, & Mahler, 1993) found that patients facing surgery were more attracted to those who were capable of meeting their desire for information (i.e., post-operative patients), compared with those who were less capable (i.e., pre-operative patients), even though the latter were more “similar” in their current state of anxiety.

Second, Roře’s (1984) utility affiliation theory posited that the tendency to affiliate under threat was determined by the perceived benefit versus cost to one’s ability to cope with the anxiety. In a test of the model, when the affiliative desires of anxious men were undercut by informing them that their anxiety about upcoming shocks would be disclosed to others, men reduced their affiliative tendencies because of “greater apprehension regarding public disclosure of their fear” (Roře & Lewin, 1988, p. 9; see also Roře, Lewin, & Hoffman, 1987; Roře, 2006). In this case, the public nature of the disclosure changed participants’ desire for the situation from information and/or comfort to maintenance of their social image, an image that would be compromised by the revelation to others that he was scared of the impending situation. Since affiliation could not facilitate his goal (image maintenance, in this case), affiliation did not result.

Third, Rabbie (1961, 1963) found that one could change the anxiety-affiliation link by either augmenting or dispelling participants’ goal of understanding their anxious reaction. Rabbie did so by generating ambiguity in some participants (i.e., “one in four find these shocks painful”) and not in others (i.e., “all find these shocks painful”) and then measuring their affiliative desires. Rabbie found that when participants had the goal of understanding their emotional state, they increased their affiliation with others. Facilitating the goal led to a stronger link between anxiety and affiliation; dispelling the goal eradicated the link. This effect has also been observed by others (Darley & Aronson, 1966; Gerard, 1963; but see Brehm & Behar, 1966).

Fourth, in a set of studies that were once thought to question Schachter’s conclusions, Sarnoff and Zimbardo (1961) agreed with Schachter (1959) that fear led to increased affiliative tendencies but argued that anxiety led to reduced affiliative tendencies. In their research, Sarnoff and Zimbardo manipulated fear using Schachter’s shock paradigm and
manipulated “anxiety” by informing male participants they would be asked to publicly suck on various objects associated with infantile “oral libido,” including an oversized nipple, a baby bottle, and a pacifier. As might be expected, fear increased affiliative tendencies, but their manipulation of anxiety reduced affiliative tendencies. This finding was demonstrated repeatedly (Dabbs & Helmreich, 1972; Firestone, Kaplan, & Russell, 1973; Lynch, Watts, Galloway, & Tryphonopoulos, 1973). However, later studies critiqued Sarnoff and Zimbardo’s conclusions by noting that Sarnoff and Zimbardo actually manipulated embarrassment, not anxiety (Paz & Amir, 1974), an emotion that involves social-image threat (like Rofé’s [1984] work reviewed previously). Indeed, later research treated Sarnoff and Zimbardo’s method as a manipulation of embarrassment that reduced affiliative tendencies (e.g., Buck & Parke, 1972; Fish, Karabenick, & Heath, 1978; M. Teichman, 1977; Y. Teichman, 1973), because individuals often experience avoidant and escapist tendencies when experiencing embarrassment (Leary, Landel, & Patton, 1996).

Of course, the two-dimensional model makes no claim about whether information or comfort is more influential in the anxiety–affiliation link. What is most important is that each study reviewed here indicates that affiliation (or not) with another person in the face of fearful stimuli is a function of the capacity of that other to meet whatever desire is most active at the time, whether it be a goal for understanding, comfort, or image maintenance.

**Initiation studies.** A phenomenon empirically related to the anxiety–affiliation link is increased attraction following a harsh, compared with mild, initiation to a group. The most notable example of this effect is Aronson and Mills (1959), who found that a female participant’s attraction to a group was more favorable after she endured a severe, rather than a mild, initiation to get into it. A meta-analysis of seven initiation studies revealed that severity of the initiation was significantly related to the liking for the group and group discussion (although the effect was small; Mullen et al., 1990).

Although cognitive dissonance is the widely accepted explanation for the effect, it is not without its critics (e.g., Finer, Hautaluoma, & Bloom, 1980; Hautaluoma & Spungin, 1974; Hautaluoma, Enge, Mitchell, & Rittwager, 1991). For instance, researchers (e.g., Lodewijkx & Syroit, 1997; 2001) have noted that among the initiation studies, only Schopler and Bateson (1962) assessed the arousal ostensibly produced by dissonance. Furthermore, field studies investigating the influence of dissonance on affiliation failed to find support for the influence of dissonance (Lodewijkx & Syroit, 1997). This lack of empirical support suggests that other or additional processes may be at play.

As with the previously discussed studies, the two-dimensional model emphasizes the notion that individuals should experience attraction to those they see as capable of facilitating their interests. As such, increased attraction toward a group after experiencing a harsh initiation would result from the impact that such an initiation has on affiliation desires and one’s assessment of the group, or group members, as capable of fulfilling such desires.

Researchers have proposed a similar explanation for the effect of initiation severity on liking for a group. Lodewijkx and Syroit (1997) proposed a model whereby a severe initiation aroused uncertainty about how one was feeling, which then led to affiliation desires to resolve that uncertainty, which then resulted in more attraction. In a laboratory study investigating this notion, these researchers failed to find an effect for negative affective state (a result inconsistent with a dissonance explanation), but found that the increased desire to affiliate, and thus, attraction to the group, was generated by a perceived ability of the group members to facilitate their understanding of their feelings. In a later study, they concluded that harsh initiations increased preferences to have affiliative exchanges with other group members, which increased their attraction to the group (Lodewijkx & Syroit, 2001). This mediational model was also supported by subsequent research that found that the anticipation of a harsh initiation increased attraction to the group via (i.e., as mediated by) participants’ affiliation desires (Lodewijkx, Van Zomeren, & Syroit, 2005).

Other research has similarly posited that the capacity of the group to provide social and emotional assistance generates the affiliative tendencies to the group. Van Duuren and di Giacomo (1997), for example, found that degrading situations (e.g., failing a test) increased affiliative motivations, because degrading situations aroused negative emotions that were expected to be reduced by the acceptance by others (see also McFarland & Ross, 1982; Kernis, Brockner, & Frankel, 1989). Similarly, Forgas (1991) found that participants who failed a test chose to pair with a socially rewarding, rather than intellectually competent, partner. That is, the affiliative tendencies of dejected participants were tied to a search for sympathy or support (see also Buck & Parke, 1972). Overall, the impact of harsh initiations on attraction to a group seems tied to the same affiliative desires that Schachter (1959) identified in his classic anxiety–affiliation studies. Harsh initiations create a desire for information or comfort that breeds affiliative tendencies to address that desire. Liking for those who facilitate the now-activated goals then results.

**Summary.** Though dissonance reduction and social comparison theories have been offered as explanations for these effects, the two-dimensional model offers a reasonable explanation for studies of attraction (a) to those who share a common fate and (b) to a group when the initiation is intense. In each case, the situation creates a novel desire that others are differentially equipped to help one fill. As discussed previously, as one’s desires change, so should one’s attraction to different others who are evaluated to be more or less capable of satisfying those desires.
However, the nature of the evidence we have reviewed only offers indirect evidence for the two-dimensional model’s perspective on these phenomena. With respect to the anticipated future distress studies, our review of 25 studies found consistent evidence that attraction was greater for those who were seen as capable of reducing participants’ level of distress. However, much of the research tested only correlational models regarding the mediating role of the capacity assessment. The same was true for the harsh initiation studies; the basic finding that affiliation and affiliative desires increase after a harsh initiation was consistent, but the data in support of a two-dimensional understanding were correlational. Such correlational findings do not allow us to make definitive claims about the role of the capacity assessment relative to the operation of other processes (e.g., dissonance).

As such, we do not suggest that dissonance or social comparison processes cannot affect attraction. The two-dimensional approach posits that dissonance (or other processes) may not be the best explanation for why attraction changes as a result of a harsh initiation, anticipation of a punishment, or other aversive stimuli. In such cases, the two-dimensional model may be a viable, and perhaps preferable, alternative.

**Similarity**

One of the most robust phenomenon in attraction literature is the “similarity effect” (Byrne, 1997): Increased similarity to a target—with respect to attitudes, personality traits, or a number of other attributes—is associated with increased attraction to the target. The similarity effect has been observed in a multitude of different populations and age groups (e.g., Byrne et al., 1971; Tan & Singh, 1995), and a meta-analysis of over 300 similarity studies observed that similarity produces a positive, moderately strong, effect on attraction (Montoya, Horton, & Kirchner, 2008). However, explanation for the link between similarity and attraction has remained a topic of debate. Whereas Byrne and colleagues have emphasized affective processes, other researchers have emphasized processes consistent with the willingness-to-benefit and capacity assessments. In the following section, we discuss three possible processes: affective, capacity-related, and willingness-related.

**Affective processes.** Byrne and colleagues (e.g., Byrne, 1971; Byrne & Clore, 1967) explain the similarity effect by proposing that a target who possesses similar attitudes is reinforcing because the target’s attributes confirm the legitimacy and accuracy of one’s own attitudes. This consensual validation, in turn, satisfies the **effectance motive**, the basic need for a consistent, logical, and accurate interpretation of the world. Dissimilar others, in contrast, offer no such consensual validation and create need-threatening inconsistency and uncertainty regarding the self. In short, Byrne argued for a behavioral reinforcement model, in which similar people are paired with good feelings that lead to attraction and dissimilar people are paired with negative affect that leads to repulsion.

This perspective has empirical support. For instance, the presentation of similar attitudes sates the effectance motive (Clore & Gormly, 1974; R. Singh, 1973, 1974), which then results in attraction to similar others (Byrne & Clore, 1967). This explanation is still used frequently in similarity research (e.g., Cann, Calhoun, & Banks, 1997; Shaikh & Kanekar, 1994), and affect has been identified as a mediator of the similarity effect in political choices (Zullow, Oettingen, Peterson, & Seligman, 1988) and consumer preferences (Meoli & Feinberg, 1989).

However, recent research has revealed that the positive affect associated with similar others only partially mediates the similarity effect (R. Singh, Ng, Ong, & Lin, 2008; R. Singh, Yeo, Lin, & Tan, 2007). Indeed, during debates with other similarity researchers, Byrne has acknowledged directly (Byrne, 1997) and tacitly (Byrne, Clore, & Smeaton, 1986) that factors other than the affect generated by consensual validation contribute to the similarity effect. Next, we review two other processes that may explain the similarity effect.

**Capacity-related process.** Kaplan and Anderson (1973) posited that the similarity effect resulted from a process involving the capacity assessment. Specifically, they hypothesized that people assume that similar others are simply “better,” more-effective people. Multiple studies have found that inferences of positive attributes increased and covaried with increased similarity (e.g., Atkinson, Brady, & Casas, 1981; L. R. Good, 1975; Jellison & Mills, 1967). Furthermore, similarity does not predict attraction once one controls for the observer’s assessment of the positivity of the target’s personality traits (i.e., the capacity assessment; Lydon et al., 1988; McLaughlin, 1970), and when the valence and similarity of personality traits attributed to a target were both manipulated orthogonally, only the valence of the personality traits predicted attraction (Ajzen, 1974).

According to this perspective, similar attitudes imply positive information about the attributes of the target, and this positive information constitutes a favorable capacity assessment of the target person that, in turn, drives attraction. In this way, the positivity of the one’s inferences about a target, rather than similarity per se, plays a central role in the similarity effect. Consistent with this view, research has shown that (a) the similarity effect is greater when information about a target person’s competence is made salient (Montoya & Horton, 2004; Montoya & Horton, 2013), (b) capacity information (i.e., the positivity of the information implied by similar attitude) mediates the effect of similarity on attraction (Montoya & Horton, 2004; R. Singh, Ho, et al., 2007; R. Singh, Lin, et al., 2008), and (c) capacity assessments partially mediate the similarity effect even when controlling for willingness information (see below; R. Singh, Yeo, et al., 2007).
Willingness-related process. A complementary interpretation of the similarity effect emphasizes the willingness-to-benefit assessment. This interpretation proposes that the similarity effect is mediated by the individual’s belief that a similar other will accept or like the individual (Aronson & Worchel, 1966; McWhirter & Jecker, 1967). As discussed earlier, people believe that when someone likes them, the admirer is likely to be benevolent in the future and will potentially help them meet their future goals (Montoya & Insko, 2008). Thus, in a study in which attitude similarity and the possibility of threat or penalties were orthogonally manipulated during a prisoner’s dilemma-type game, the possibility of threats increased liking in low-similarity dyads, but it reduced liking in high-similarity dyads. The researchers concluded that the possibility of threats/penalties in the high-similarity dyads created “an air of uncertainty to an interaction that is typically based on confidence: confidence in trust, mutual support, and expectations of supportiveness, generosity, and cooperativeness” (Tornatzky & Geiwitz, 1968; p. 126). This interpretation is consistent with a willingness-to-benefit explanation for the similarity effect.

Such an explanation is bolstered by a number of empirical findings. To start, similarity, compared with dissimilarity, leads to greater expectations of acceptance by, and more liking for, a target (Inske, Sedlak, & Lipsitz, 1982; McWhirter & Jecker, 1967; see also Walster & Walster, 1963). Indeed, expectations of acceptance and similarity tended to be highly correlated (e.g., \( r = .87 \) from McWhirter & Jecker, 1967). Furthermore, when participants were asked to estimate the degree to which a stranger would like them as a function of attitude similarity with the stranger, they produced an estimate that was consistent with Byrne and Nelson’s (1965) formula for the influence of similarity on attraction (Gonzales, Davis, Loney, Lukens, & Junghans, 1983; see also Moss, Byrne, Baskett, & Sachs, 1975).

In a more sensitive test of the influence of willingness-to-benefit information on the similarity effect, Condon and Crano (1988) found that (a) the effect of similarity on attraction was negligible when perceived acceptance was included as a predictor, and (b) the effect of perceived acceptance on attraction was strong even when similarity was included as a predictor. These findings indicate that willingness information (in the form of perceived acceptance) plays a strong role in the similarity effect. Indeed, other work has found that (a) inferred attraction mediates the similarity effect (Simons, 2008; R. Singh et al., 2009); (b) inferred attraction mediates the similarity effect even when controlling for the assessment of a similar target’s capacity (R. Singh, Yeo, et al., 2007), (c) when liking is expected from another, the degree of similarity does not affect attraction (Aron, Steele, Kashdan, & Perez, 2006), and (d) individuals prefer to affiliate with dissimilar others when liking is assumed (Walster & Walster, 1963). Overall, the evidence supports the notion that similarity affects attraction because it indicates that a target is willing to accept the perceiver.

Comparing explanations for the similarity effect. Recent work has assessed the relative merit of the different explanations for the similarity effect. Most notably, research by Singh and colleagues (R. Singh, Yeo, et al., 2007; R. Singh, Ng, et al., 2008) manipulated similarity between a participant and a target, then assessed perceived acceptance (as an index of willingness-to-benefit), assessments of the target’s perceived capacity, positive affect, and attraction to the target. When the mediators were tested individually, affect partially mediated the similarity effect, while willingness and capacity assessments fully mediated the similarity effect. When included as simultaneous mediators, the similarity effect was more strongly mediated by the willingness assessment than by the capacity assessment, and affect failed to mediate the link. R. Singh, Yeo, et al. (2007) concluded that an assessment of willingness and capacity are necessary to fully account for the similarity effect. Furthermore, a meta-analysis of the similarity effect that compared Byrne’s (1971) affect-driven reinforcement model to the cognition-driven explanation for the similarity effect (i.e., information integration perspective of Kaplan & Anderson, 1973) found that the similarity effect was greater when the measurement of attraction followed, rather than preceded, the capacity measurement of the target person (Montoya & Horton, 2013). This result provides further evidence for the importance of capacity processes to the similarity effect.

Summary. There is strong support that the similarity effect is mediated by willingness-to-benefit and capacity assessments. When either capacity or willingness information is controlled, similarity no longer leads to attraction. The findings for affect are more mixed: When positive affect is included alone as a mediator, it partially mediates the similarity effect. However, tests in which affect is included simultaneously with willingness and/or capacity information reveal that affect plays no unique role in the similarity effect.

There are, of course, many other perspectives on the causes of the similarity effect, and we discuss two of the most popular relative to the two-dimensional model. First, the rewards of interaction model (D. Davis, 1981; Werner & Parmelee, 1979) posits that similarity information informs people that rewards are likely to follow (as when similar attitudes suggest that common activities/hobbies are present and that such interactions will be comfortable and without conflict), and attraction follows this expectation of future rewards. Such an explanation can be interpreted as consistent with the two-dimensional approach. That is, expectations of reward may result from an understanding that the other person is “good” (i.e., capacity assessment) because of the qualities associated with a person who would engage in “good” activities. Alternatively, such expectations might come from a belief that the other person is willing to provide beneficial interactions because the other likes the individual (i.e., willingness assessment). Second, Rosenbaum (1986) proposed that the positive relation between similarity and attraction.
resulted from the repulsion caused by the dissimilar attitudes—the greater the number of dissimilar attitudes, the greater the repulsion. However, repeated tests of Rosenbaum’s (1986) findings concluded that although dissimilar attitudes were more impactful on evaluations than similar attitudes (e.g., Jia & Singh, 2009; R. Singh & Ho, 2000; R. Singh & Teoh, 1999), the standard principles of the similarity effect still held (Byrne et al., 1986; R. Singh & Tan, 1992; Smeaton, Byrne, & Murnen, 1989; Tan & Singh, 1995). Given the evidence that the repulsion effect may result from the same processes as those operating with the similarity effect, it would be entirely consistent that the repulsion effect is also compatible with the two-dimensional model.

General Discussion

The purpose of this article was to propose a model for understanding the experience of interpersonal attraction. This model posits that the experience of attraction is strongly affected by the degree to which people perceive that another person is capable and willing to act benevolently toward them. The proposed model is necessarily broad to accommodate the diversity of findings that populate the attraction literature. Despite its breadth, however, the model remains testable and generative. Indeed, the latter half of this review considered the merit of the two-dimensional model for understanding several phenomena that involve attraction (e.g., pratfall effect, similarity effect) whose original theoretical underpinnings were far removed from the evolutionary person perception ideas that guided the formation of this model. For each reviewed phenomenon, research across a diverse set of attraction phenomena provided considerable evidence that willingness and capacity assessments were tied to the experience of interpersonal attraction. In each case, the two-dimensional approach made specific predictions that willingness and capacity assessments regulate the experience of attraction, and in each case, evidence was consistent with such theorizing. Of course, we did not address the model’s fit with every conceivable finding in the literature; however, we attempted to be thorough and objective in our consideration of the most robust and productive programs of research.

Comparison to Other Two-Dimensional Approaches

The two-dimensional model can be understood relative to previous models with a two-dimensional structure. To begin, the model expands on extant two-dimensional person perception approaches in two distinct ways. First, extant approaches do not consider the collective impact or relation between capacity and willingness-to-benefit assessments. Indeed, such models identify the two assessment dimensions without consideration of the impact of the dimensions on attraction (or other phenomena). Furthermore, these approaches do not account effectively for phenomena affected by reduced willingness due to a high capacity assessment. More specifically, they do not explain effectively the standard pratfall effect, whereby attraction to a high-quality other is reduced because rejection is perceived as likely. The advantage of the proposed two-dimensional model is that it explains changes in attraction due to the influence of a deflated willingness assessment. Second, person perception approaches do not consider, as the presented two-dimensional model does, the goals of the perceiver and how such goals influence the importance and valence of judgmental dimensions for feelings about and behavior toward a target. As noted previously, the judgments one makes of a target are affected by activated goals, and affective and behavioral responses flow from such judgments. The two-dimensional model describes the motivational component of interpersonal attraction.

In addition, the two-dimensional model can be differentiated from previous two-dimensional approaches by the proposed relation between willingness and capacity. That is, the two-dimensional model proposes that a weighted combination of the willingness and capacity assessments determine attraction, but most two-dimensional approaches either posit a multiplicative relation (e.g., Wojciszke, Abele, & Barylka, 2009) or do not specify the relation. In addition, the relations among willingness, capacity, and attraction may be curvilinear rather than linear, as has been posited by older expectancy-value models (Fishbein, 1967; Fishbein & Ajzen, 1975), but has not been embraced by recent person perception/relationship models. In fact, there is evidence for such curvilinear effects. Specifically, consistent with the economic notion of “diminishing marginal utility” (Bernoulli, 1954), changes in low levels of willingness may be more influential to attraction than similar changes at high levels of willingness (VanLear & Trujillo, 1986). Research on uncertainty, for example, indicates that the possibility of acceptance (a willingness manipulation) can produce a strong behavioral attraction response (e.g., Whitchurch, Wilson, & Gilbert, 2011), but the “bump” for attraction may not hold for higher levels of willingness (Larzelere & Huston, 1980).

Finally, the proposed two-dimensional model makes specific predictions regarding conditions in which affective attraction should differ from behavioral attraction—and when the level of affective and/or behavioral attraction would differ from the nature or valence of the capacity or willingness-to-benefit assessments. Next, we discuss two such predictions.

Divergence affective/behavior from willingness/capacity. As discussed earlier, increased capacity can reduce the favorability of the willingness assessment, which results in reduced behavioral attraction, thus producing a divergence between the level of behavioral attraction and the valence of the capacity information. This difference has been observed repeatedly (e.g., Herbst et al., 2003).
Divergence of affective from behavioral attraction. Research indicates that affective and behavioral attraction most commonly diverge because of (a) information seeking and (b) self-interest considerations. With respect to the former, information seeking or social comparison processes may motivate more behavioral attraction than affective attraction to help people determine how they should evaluate others. Berger and Calabrese (1975), for example, proposed that uncertainty regarding how another feels about us would produce affiliative behavior to reduce such uncertainty, but such uncertainty would also be paired with low levels of affective attraction. With respect to self-interest considerations, such considerations may reduce behavioral attraction to protect the person from negative feedback. Research from the matching and pratfall literature supports such a contention. A second example of the difference between affective and behavioral attraction comes from the dating literature, which has observed differences between one’s preferences for physically attractive others (i.e., affective attraction reported in non-threatening contexts) and selection of a date whose physical attractiveness matches one’s own (behavioral attraction responses reported when rejection is feared from attractive people; for example, Kurzban & Weeden, 2005; Todd, Penke, Fasolo, & Lenton, 2007).

Limit of the Two-Dimensional Model

An important unanswered question is the extent to which the two-dimensional model effectively describes the panoply of attraction phenomena. Although we have devoted substantial attention to interpreting attraction phenomena in light of a two-dimensional framework, each of the discussed phenomena reflect rather controlled and “explicit” attraction assessments that were reported only after conscious deliberation. The model’s emphasis on willingness-to-benefit and capacity assessments fits well with such explicit experiences of attraction. However, although not established to this point, our model may also prove useful for understanding “implicit,” non-conscious attraction phenomena. Assuming that willingness and capacity assessments can occur rapidly and automatically (e.g., Wentura, Rothermund, & Bak, 2000; Willis & Todorov, 2006), the model may also be able to describe implicit, automatic effects, including the mere exposure effect (Zajonc, 1968) or implicit egotism (Pelham et al., 2005), and other effects that are hypothesized to proceed as a result of automatic processes.

As an example, the mere exposure effect (Zajonc, 1968) describes the increase in liking that results from repeated exposure of non-valenced stimuli (for a review, see Bornstein, 1989). Researchers have argued this effect results from an automatic affective process that operates without cognitive processes (Zajonc, 2001) or with minimal cognitive processing (Bornstein & D’Agostino, 1992). Alternatively, the two-dimensional model would posit that the mere exposure effect, as it applies to evaluations of inanimate stimuli (e.g., ideographs, polygons), results from the non-conscious impact of repetition of a stimulus on one’s perception that a stimulus is “correct” or “how it should be”—an explanation consistent with a simple manipulation of the capacity assessment. Put another way, this two-dimensional perspective capitalizes on the notion that people prefer prototypes because prototypes are “correct,” a notion that is consistent with arguments posited by Mandler (1984), who proposed that individuals experience positive affect from stimuli congruent with an established schema because it is “as it should be” (p. 207).

There is reason to believe in such an interpretation: (a) repeated exposure to a stimulus increases assessments of the “prototypicality” of a stimulus (Farkas, 2002; Martindale & Moore, 1988), (b) judgments of the prototypicality of faces are strongly related to how attractive those faces are judged to be (D. Jones & Hill, 1993; Light, Hollander, & Kayra-Stuart, 1981), possibly because they are evaluated as “healthy” (i.e., evolutionarily “capable”) and as how they should be (Thornhill & Gangestad, 1993; see also Symons, 1979), and (c) assessments of prototypicality mediate the attraction that results from repeated exposure (Winkielman, Halberstadt, Fazendeiro, & Catty, 2006, Experiments 1 and 2; Zebrowitz, White, & Wieneke, 2008). In this way, evidence indicates that prototypicality mediates and facilitates the expression of the mere exposure effect.

As a second example, consider implicit egotism. Implicit egotism refers to individuals’ preference for objects, names, professions, and brands associated with the self (Pelham et al., 2005). People are particularly attracted to their first and last initials and their birthdates and to people, professions, and places that share their letters or numbers (J. T. Jones, Pelham, Carvallo, & Mirenberg, 2004; Pelham, Mirenberg, & Jones, 2002). From the two-dimensional perspective, implicit egotism may reflect manipulations of the willingness-to-benefit assessment, such that self-similarity serves as a cue to trust. In J. T. Jones et al. (2004; Study 5b), for example, participants learned that their partner’s arbitrary experimental code number (e.g., 10-26) was either the same as their birthday or was different. As was expected, participants experienced greater behavioral attraction for those whose code number matched their birthday. However, when “anticipated liking” (i.e., an assessment consistent with willingness) was included in the model as a mediator, code number similarity no longer predicted attraction, indicating that the willingness assessment may be important to understanding the effects of implicit egotism.

In sum, the two-dimensional model identifies processes that may play an important role in creating “implicit” attraction effects. The model makes specific predictions regarding the regulation of the mechanisms underlying the mere exposure and implicit egotism effects (and, by extension, other associative effects, such as behavioral mimicry, Chartrand & Bargh, 1996; or evaluative conditioning, De Houwer, Thomas, & Baeyens, 2001; Hofmann, De Houwer, Perugini,
1. Notes

Funding to the research, authorship, and/or publication of this article.

Declaration of Conflicting Interest

The two-dimensional model of attraction indicates that the degree of experienced attraction toward another person can be determined by the following question, “How ‘good’ is he or she and does he or she like you?” The first part of the question reflects the capacity assessment, which considers the overall aptitude and competence of the other person on goal-relevant dimensions. The second part relates to the willingness-to-benefit assessment, which is commonly inferred from the assessment of perceived acceptance. This two-dimensional approach is based on the fundamental dimensions outlined by the person perception literature, and we have provided evidence for the model by investigating it relative to the dominant attraction phenomena—including the reciprocity of liking, the matching hypothesis, and the similarity effect. In each case, we found strong evidence that changes in willingness and capacity assessments were associated with changes in attraction.

As noted at the outset of this article, the progress of research on interpersonal attraction has been delayed by differences between theoretical models, inconsistencies in defining attraction, and disagreements regarding which methods are best suited to the study of attraction. The proposed model stands to move the field forward by (a) providing a comprehensive model that is broadly applicable to experiences of attraction, (b) differentiating between and making separate predictions for affective and behavioral attraction, and (c) emphasizing the role of experimental research in illuminating the processes that underlie the experience of interpersonal attraction. Although we cannot completely exclude other processes from consideration, we conclude that the existing attraction research supports a two-dimensional approach to understanding interpersonal attraction phenomena.

References


2. It is also worth noting that a two-dimensional perspective has also been applied to personality traits. For example, whereas personality is commonly viewed using the Big Five (McCrae & Costa, 1989), theorists have reinterpreted the typical Big Five solutions in terms of “superfactors” that resemble capacity and willingness dimensions. Wiggins (1991), for example, argued the Big Five judgments constituted assessments of dominance/ambition (akin to the capacity assessment) and nurturance/warmth (akin to the willingness assessment). Similarly, Digman (1997; see also Buss, 1989b) concluded that the Big Five dimensions loaded onto more general dimensions of personal growth (dominance/agency) and socialization (nurturance/warmth), factors that we would regard as consistent with the capacity and willingness assessments. More recently, Abele and Wojciszke (2007) concluded that Bakan’s (1966) agency versus communion dichotomy mapped onto dominance/competence (capacity) and morality/warmth (willingness) dimensions.

3. The theoretical connection between our definition of capacity and the motivated cognition literature (e.g., Bargh, 2005) indicates that conceptual differences can be specified between the present model and reinforcement-affect model’s (e.g., Byrne & Clore, 1970) view of attraction as resulting from the target’s past association with (or ability to be) a source of reward. Specifically, whereas the two-dimensional model considers an individual’s internal and subjective goals as critical to predicting attraction responses, reinforcement models focus on externally realized goals as important (for a more complete discussion, see Bargh, Gollwitzer, & Oettingen, 2010; Finkel & Eastwick, in press). Second, and in contrasts to a reinforcement approach, the two-dimensional model posits that unconsciously activated goals would predict attraction response (e.g., Chartrand & Bargh, 2002).

4. Further evidence for the different influences of physical attractiveness on capacity and willingness-to-benefit assessments may be seen in the physical attractiveness stereotype, in which attractive others are judged more favorably than less-attractive others on capacity dimension (social competence, adjustment, potency, intellectual competence), but not on willingness dimensions (integrity, concern for others; Eagly et al., 1991).


Jones, A. P., James, L. R., & Bruni, J. R. (1975). Perceived leadership behavior and employee confidence in the leader as moderated by job involvement. *Journal of Applied Psychology*, 60, 146-149.


71, 54-70.

Journal of Personality and Social Psychology, 12, 186-204.

9, 327-349.

28, 379-388.

291-296.

270-296.

287-318.

271-296.

279-309.

271-296.

279-309.


