THE SECOND FACE OF TRUST:  
REFLECTIONS ON THE DARK SIDE OF  
INTERPERSONAL TRUST IN ORGANIZATIONS  

Daniel J. McAllister

Every truth has two sides. It is well to look at both sides before we commit ourselves to either.  
Aesop

Oftentimes, to win us to our harm, the instruments of darkness tell us truths, win us with honest trifles, to betray's in deepest consequence  
(Macbeth, Act I, Sc. 3, line 123).

Management scholars and professionals agree on the importance of interpersonal trust relationships in organizations—trust is good, desirable, even essential, for organizations to function properly (Shaw, 1997; Kramer & Tyler, 1996; Arrow, 1974; Thompson, 1967). For organizations, trust relationships can enhance levels of coordination and efficiency (Bradach & Eccles, 1989; Pennings & Woiceshyn, 1987), reduce the need for monitoring and controls (Ouchi, 1979; Pennings & Woiceshyn, 1987), and increase levels of interpersonal helping and coordination-enhancing behavior (Organ, 1988; McAllister, 1995). For organization members,
trust relationships enhance the quality of life, providing needed support, pleasure, meaning, and purpose (Baumeister & Leary, 1995; Stevens & Fiske, 1995; Lobel, Quinn, St. Clair, & Warfield, 1994).

In our present enthusiasm for and interest in the richness of trust relationships and their positive implications for organizations, we have given little attention to the limits of trust (Barber, 1983), the potential unintended consequences of trust, and the implications of trust’s limitations for organizations and their members. Trust relationships can expose organizations and their members to risks of betrayal under conditions that augur against deception detection (Shapiro, 1987, 1990). Much of the white-collar crime that takes place in organizations is facilitated by the existence of trust relationships (Albrecht, Wernz, & Williams, 1995; Friedrichs, 1996; Shapiro, 1990). Beyond the losses stemming from betrayals of trust, emergent trust relationships can serve interests other than those of the organizations in which they exist. Collective connivances involve considerable trust (trust or honor among thieves), and they can yield considerable damage (Granovetter, 1985; Wintrobe & Breton, 1980). A balanced perspective on interpersonal trust and the social processes underlying trust’s development and maintenance is needed—a perspective that acknowledges the limitations and possible negative consequences of trust as well as the benefits, and addresses the means by which the potential pitfalls of trust are managed.

As a first step towards balance in our understanding of the nature and functioning of trust relationships in organizations, I will attempt to develop a profile of the second face of interpersonal trust—trust’s dark side—one in which the problematic qualities and dysfunctions of trust relationships receive acknowledgement, one in which the generic prescription that “more trust is better” certainly does not apply. The focus will be the conditions under which (1) trust relationships are manipulated and confidences are betrayed, (2) trustworthiness is misattributed, (3) failing and flawed trust relationships persist, and (4) substantial personal and organizational costs accompany the termination of trust relationships. Discussion begins with consideration of the nature of interpersonal trust relationships in organizations, continues with consideration of the dark side or potentially unintended consequences of trust relationships in organizations and the social consequences leading to trust failure, and concludes with consideration of practical managerial implications and research questions that follow from the analysis.

UNDERSTANDING INTERPERSONAL TRUST

Defining Interpersonal Trust

At a basic level, interpersonal trust can be defined as an individual’s belief in, and willingness to act on the basis of, the words, actions, and decisions of another (McAllister, 1995). From this perspective, trust is intimately concerned with the
confident in others that an individual brings to risky interdependent situations. Confidence can be expressed in a number of ways including a belief in the predictability and reliability of others (Mayer, Davis, & Schoorman, 1985), a belief in the benevolent motives of others (believing that others take partner interests into account when they make decisions) (Deutsch, 1960), and a willingness to act on the basis of these beliefs (Axelrod, 1984). As such, effective trust relationships are characterized by secure interpersonal attachments and provide the foundations for productive exploration and learning, as well as broader affiliation (Kahn & Kram, 1994; Rothbard & Shaver, 1994; Hazan & Shaver, 1990).

An important element in this definition of trust is the acknowledgment that trust extends beyond beliefs about the predictable behavior of others (Mayer, et al., 1995) to address their motives or goodwill when making decisions (Lewicki & Bunker, 1995). Following Simon (1957), decision premises provide a more fundamental basis for trust than compliance with specific behavioral expectations. Indeed, leaders, as decision makers, frequently have to make decisions that appear to go against the interests of others (laying off associates, reducing budgets, and such). Under such conditions, however, the trust of others may still persist insofar as it is grounded in beliefs about motives rather than the specific undesired behaviors (Tyler & Degoey, 1996; Read, 1962).

A further distinction can be made between fundamental forms of trust by separating the rational bases for trust from the socio-emotional factors grounding trust decisions (Lewis & Wiegert, 1985). For instance, Ring (1996) distinguishes fragile trust (trust that permits economic actors to deal with each other, but only in guarded ways) from resilient trust (trust that rests...in the predictability of the goodwill of others). Likewise, Lewicki and Bunker (1995) distinguish between knowledge-based trust and identification-based trust. For Lewicki and Bunker (1995), knowledge-based trust entails “knowing the other sufficiently well so that the other’s behavior is anticipatable” (1995, p. 12) and identification-based trust entails “full internalization of the others’ desires and intentions” (1995, p. 13). Within the present discussion we will frame this fundamental distinction as one between cognition-based trust (where trust is grounded in an individual’s dependability, reliability, and professionalism) and affect-based trust (where trust is grounded in the emotional bonds connecting interdependent individuals to one another in a relationship) (McAllister, 1995).

Cognition-Based Trust

For cognition-based trust, confidence is grounded in an individual’s track-record and reputation for dependability, reliability, and professionalism. Role expectations and professional norms circumscribe the domains within which people’s words, actions, and decisions, can be provisionally trusted (Zucker, 1986; Shapiro, 1990). Provisional trust derived from impersonal standards (credentials, certification, and such) is transformed into personal cognition-based trust through
interaction over time. Shared understandings and expectations (value congruence), the product of common socialization experiences, facilitate the emergence of trust by reducing the likelihood of conflicting or ambiguous signs of trustworthiness (Sitkin & Roth, 1993). Once established, and in the absence of evidence bringing trustworthiness into question, cognition-based trust persists (Garfinkel, 1973; Gambetta, 1988).

Even the best of evidence cannot rule out the possibility that trust is misplaced. Doctors and professionals occasionally do provide medical and professional services without proper credentials and betray the confidences of their patients (Gabbard & Nadelson, 1995), lawyers deceive and lie to their clients (Lerman, 1990), dependable peers at work still let others down from time to time, and damage is done. Nevertheless, available information does present a rational foundation for trusting—for taking steps of faith to behave with certainty in the face of uncertainty.

Affect-Based Trust

Affect-based trust, likewise, incorporates an individual's belief in, and willingness to act on the basis of the words, actions, and decisions of others. As people make emotional investments in personal relationships, express genuine care and concern for the welfare of others, and come to understand these sentiments as being reciprocated, confidence in partner words, actions, and decisions increases. Affective foundations do not rule out the possibility of betrayal—infidelities still occur, and personal confidences are betrayed from time to time. Yet affective foundations, like those that are cognition-based, enable trusting individuals to make trusting leaps of faith, behaving with certainty in the face of uncertainty.

The fact that affect-based trust and cognition-based trust exist as distinct forms of trust does not imply that affect-based trust is somehow non-rational and/or that affective factors do not influence cognition-based trust decisions. Acknowledgment that a relationship of care and concern exists, and that both parties value each other and the relationship itself, should provide, in part, a rational basis for expecting partners to behave in ways that do not harm each other, and that sustain relationship. Affect-based trust is typically vested in others only after a certain level of cognition-based trust is established. Baseline confidence in partner dependability and reliability precedes the development of more in-depth affiliations (Lewicki & Bunker, 1995; McAllister, 1995).

Developmental Processes Underlying Interpersonal Trust

A general framework for understanding the social processes underlying the development of cognition- and affect-based trust can be outlined, together with specification of the assessment criteria for trustworthiness evaluations. Social construction (Zucker, 1986; Shapiro, 1990) and social exchange (Blau, 1963) pro-
cesses underly the development of cognition-based trust, and norms of equity and equality provide acceptable standards for assessing trustworthiness from a cognition-based perspective (Deutsch, 1975). Social exchange considerations give way to attribution concerns as affect based trust develops (Holmes & Rempel, 1989; Mills & Clark, 1994). Norms for need-based responding provide foundations for affect-based trust assessments (Clark & Chrisman, 1994).

Foundations for Cognition-Based Trust

Given the complexity of modern society, it is amazing how much social order we can take for granted (Luhmann, 1979; Barber, 1983). Within our organizational society, social institutions including laws, professions, occupational categories, and such provide advanced foundations for both impersonal and personal professional relationships (Zucker, 1986; Shapiro, 1987, 1990). Formal role descriptors (engineer, doctor, mechanic) and related credentialling designations (professional engineer, medical doctor, tradesman) provide a partial foundation for trust. Provided that professionals did not receive their credentials by force and fraud, we can expect them to be dependable and reliable within the confines of their role descriptions (Shapiro, 1987). These foundations for trust are most salient in the early stages of trust development, when the social fabric needed for richer trust development is yet to be woven.

With familiarity, it becomes increasingly possible for individuals to make confident assessments of the dependability and reliability of others (Luhmann, 1979)—provisional assessments of trustworthiness stemming from institutional arrangements gain an experiential base. With repeated interaction it becomes possible for norms of fair exchange and reciprocity to be established. The scope of goods and services given and received in social exchange and the time-frame prescribed for transaction completion can expand tremendously. Mutually beneficial exchange, characterized by equity and/or equality (Deutsch, 1975; Mannix, Neale, & Northcraft, 1995) provides a respectable basis for the development of cognition-based trust.

Foundations for Affect-Based Trust

Attributional processes that illuminate the motives underlying the behaviors of relationship associates govern the development of affect-based trust. In close relationships, observe Holmes and Rempel (1989), “people’s expectations relevant to trust center on their perceptions of their partner’s attitude toward the relationship, on the perceived quality and intensity of the affective bond…. Expectations that form the very core of trust are those that focus on the partner’s responsiveness to needs in situations of dependency” (1989, p. 188). From this perspective, affect-based trust attributions develop as partner macromotives of care and concern are discerned (Holmes, 1978), as when individuals show themselves to be sacrificially
responsive to partner needs. Behavior recognized as personally chosen rather than role prescribed, serving to meet legitimate needs, and demonstrating interpersonal care and concern rather than enlightened self-interest is critical to the development of affect-based trust (McAllister, 1995).

The logic underlying the development of affect-based trust, therefore, differs considerably from that of cognition-based trust. This distinction parallels Mills and Clark's distinction between the logics of exchange and communal relationships (Clark & Mills, 1979). Mills and Clark (1994) summarize key research findings on the difference between these two forms of relationships:

We have found that when exchange relationships are desired, people react positively to receiving immediate compensation for favors and positively to requests for repayment of favors they have requested from others; they keep track of individual inputs into tasks for which there will be a joint reward and they feel exploited when they give help that is not specifically repaid. In contrast, when subjects are led to desire communal relationships, they react negatively to receiving immediate repayment for favors; they do not appear to keep track of individual inputs into joint tasks, and they do not feel exploited when their help is not specifically repaid. Instead, subjects desiring communal relationships with others are more likely than those desiring exchange relationships to keep track of the others' needs even when they cannot help, to help others, to respond to the others' sadness with increased helping, to feel good about having provided help to the others relative to not being able to help, and to welcome expressions of emotion from the other (1994, p. 32, references omitted).

Affect-based trust relationships are best viewed as being intendedly communal in nature. Responsiveness to partner needs provides the appropriate standard for assessments of affect-based trust (Clark & Chrisman, 1994; Folger, Sheppard, & Buttram, 1995; Deutsch, 1975). Indeed, following Mills and Clark (1994), it is precisely because behavior in affect-based trust relationships runs counter to alternative justice standards—equity and equality and fair exchange—that clear attributions regarding partner motivations are possible.

The Relationship Between Affect-and Cognition-Based Trust

Consistent with both the social-psychological literature on trust in close relationships (Homes & Rempel, 1989; Rempel et al., 1985) and recent conceptual work on trust within the organization sciences (Lewicki & Bunker, 1995), affect-based trust is built upon a foundation of cognition-based trust. People's baseline expectations for dependability and reliability must be met before clear attributions concerning peer motives are possible and before deeper personal investments in trust relationships take place.

This serial ordering of cognition- and affect-based trust has practical significance. First, conditions giving rise to the need for greater trust promote the emergence of affect-based trust in particular. As the social contexts of working relationships become more hostile and unpredictable, interdependent individuals must do much more than just depend on each other to do their jobs dependably and
reliably—mutual assistance and support becomes critical and beliefs about partner priorities and motives take on added significance. Personal commitments to one another become an increasingly important factor for high performance collaboration (Katzenbach & Smith, 1994). Not surprisingly, cognition-based trust alone becomes insufficient, and affect-based trust emerges as central. Second, conditions that lead individuals to value and depend upon trusted partners promote the emergence of affect-based trust. At the outset of working relationships, dependable, reliable, and professional peers are extremely desirable. While increased cognition-based trust directed towards a proven peer may suggest reduced levels of insecurity, the prospects of losing such a colleague may give rise to a heightened sense of felt insecurity in the relationship (Holmes, 1991). Under this condition, a shift in focus of attention from whether the colleague is reliable and dependable to whether he or she cares about my best interests, and will continue to do so in the future, seems likely.

We have developed a theoretical framework for understanding the nature and development of interpersonal trust relationships in organizations. Again, much of what we know about trust in organizations tells us that trust is essential and beneficial (Bradach & Eccles, 1989; Mayer, et al., 1995; Kramer & Tyler, 1996), reducing the need for monitoring and controls (Ouchi, 1979; Pennings & Woesheyn, 1987) and increasing levels of interpersonal helping and coordination-enhancing behavior (McAllister, 1995). Both cognition- and affect-based trust are prevalent forms of trust in organizations, and the naive presumption may well be that this trust is serving its understood purpose—promoting efficiency and effectiveness. The focus of the remaining analysis will be on exploration of limitations in this understanding and the ways in which expectations concerning the benefits of trust may go unfulfilled.

THE DARK SIDE OF INTERPERSONAL TRUST: REASON FOR CAUTION

Trust Failures: Misattribution and Betrayal

The potential benefits of interpersonal trust relationships notwithstanding, management scholars have yet to consider the risks of these relationships. To the extent that trust relationships leave individuals and organizations open to betrayal and abuse, a singular emphasis on the efficiencies of trust may do more harm than good (Barber, 1983). In the present analysis the focus is on four elements of the dark side of trust: the susceptibility of trust relationships to manipulation and abuse, risks of misattributions of trustworthiness, the tendency towards persistence in failed and failing trust relationships, and the individual and organizational carnage/loss associated with the breakdown of trust relationships.

First, while management scholars have focused on the organizational efficiencies of trust, sociologists have documented trust's critical role in white-collar
crime (Sutherland, 1949; Cressy, 1953; Blum, 1972; Shapiro, 1987, 1990; Friedrichs, 1996). For these scholars, "because trust norms...disarm the protections that wary principals would otherwise erect around social and economic exchange, they increase the opportunities for abuse of fiduciary relationships at the same time that they make these relationships possible" (Shapiro, 1990, p. 350). Recent examples highlighted in the popular press—Joseph Waldholz's ability to orchestrate a $1.7 million check-kiting scheme ("Tearful Waldholtz," 1995); Gary Fairchild's embezzlement of funds from Winston & Strawn, a 455-lawyer firm, over a period of seven years (Dillon, 1994)—reveal the protections trust provides for those who would engage in craft and deception. Albrecht, Wernz, and Williams (1995), advising practitioners on managing fraud in organizations, observe that it is generally people we "trust" and "have confidence in" who can and do commit frauds (1995, p. 267). The largest group of fraud perpetrators by age consists of people between 36 and 45 years of age, a group comprised primarily of middle managers who have worked themselves into positions of trust.

The second dark side of trust is revealed where we see trust misattributed and failures in trust go unrecognized. For example, alcoholism, drug abuse, and mental illness among managers and professionals can have devastating consequences for organizations, but frequently go undetected and untreated (Speller, 1989). Conditions of impairment may significantly limit the ability of individuals to effectively perform their jobs—to use time effectively, collaborate with others, and make wise professional and business decisions. In addition, executive impairment due to substance abuse may result in other abuses of trust (embezzlement, theft from clients, and such) (Albrecht, et al., 1995).

In the terminology of applied statistics, misattributions of trust represent Type II errors—errors of trusting where distrust is appropriate. Why does misattribution take place? Why is untrustworthiness associated with executive impairment, occupational stress, substance abuse, and such not detected and addressed? One important ingredient in the puzzle is trust—trust acquired in the process of moving up in an organization's hierarchy, trust accorded to individuals filling executive and professional positions (characterized by considerable autonomy and a sizable time-span of discretion). Additionally, the absence of evidence that brings trustworthiness beliefs into question for extended periods of time—the fact that impaired executives may function "normally" most of the time—may serve to mask the gravity of the situation. Unfortunately, it may be only after irreversible damage is done, with the benefit of hind-sight, that the true limitations in an impaired executive's trustworthiness are truly recognized (Baron, 1993).

Third, the losses associated with manipulated, abused, and misplaced trust are compounded as such situations persist, warning signs of untrustworthiness are not heeded, and evidence of untrustworthiness remains unacknowledged. It is discouragingly remarkable to see how much time can pass by before impaired employees are assisted through employee assistance programs (EAPs) (Speller, 1989), how long warning signs of employee instability are usually visible before extreme
workplace violence takes place and it is too late (Baron, 1993). and the duration of
time that can pass before patterns of embezzlement and employee fraud are
detected (Albrecht, et al., 1995). Indeed, a good portion of this behavior ultimately
goes undetected and unaddressed (Greenberg, this volume). For Albrecht and
associates, the core problem is that the employees best able to detect craft and
fraud are usually the least prepared to expect and detect corruption, and the oppor-
tunity structures and reward systems in place augur against timely reporting. In the
present analysis, considerable attention will be given to the perceptual and attribu-
tional processes governing affect-based trust that guide the labeling of behavior
and its trust implications.

Fourth and finally, the dark side of trust is revealed in the aftermath of
betrayal—what happens when trust is betrayed, the betrayal is perceived and
understood as such, and consequences follow. By its very definition, betrayal sig-
nifies a violation of the trust expectations on which personal relationships are
based (Jones & Burdette, 1994). In ongoing working relationships betrayal can
take many forms and vary in terms of its perceived seriousness. While betrayals
that involve sexually coercive and harassing behavior may be most serious and
reacted to most strongly (Yeatts, 1994), other forms of personal betrayal—coer-
cive or threatening behavior, withholding promised support, acts that reduce
employee opportunities, blaming employees for personal mistakes, unfair com-
ensation practices, favoritism, improper dismissal, over-reacting to problems or
situations, and misuse of or failure to protect private information—are also salient
to employees.

The negative consequences of trust betrayals can be extreme and lasting.
Research evidence from the medical and psychological/psychiatric professions
indicates that professional-client sexual intimacy represents a fundamental breach
or betrayal of trust (Gabbard & Nadelson, 1995; Sherman, 1993). For betrayed cli-
ents, the consequences may include mental instability that limits the individual's
ability to work effectively, reduced self esteem and self confidence, increased
anger, and crippled ability to trust others (Sonne, Meyer, Borys, & Marshall 1985).
Research findings indicate that the dynamics of betrayal in therapist-client rela-
tions generalize to various professional fiduciary relationships in general, includ-
ing lawyer-client and professor-student relationships. Generalizability to
manager-employee relations seems likely. Above and beyond betrayals that follow
from sexual indiscretion, betrayals of trust in general, and betrayals of affect-
based trust in particular, are likely to negatively impact the betrayed person's reg-
ular job performance, her or his self concept, and her or his ability to trust others.
Research findings reported by Jones and Burdette (1994) and Hansson, Jones, and
Fletcher (1990) show that betrayals of trust relationships by coworkers are preva-
 lent, that memories of betrayal episodes are harbored for decades, and that the
overwhelming effects of betrayal episodes are negative.

Overall, it appears that there is a certain dark side to interpersonal trust. Trust in
the competencies and motives of others, when that trust is misplaced, can have
damaging consequences. The following section explores the factors that contribute to the fragility of trust relationships and the susceptibility of trust to betrayal and abuse.

Misattributions Of Trustworthiness: A Framework For Understanding Trust Failures

Interpersonal trust develops as the product of interpersonal interaction within a social setting. From a dramaturgical perspective, the nature of the setting, the processes of social interaction, attributes of the trusting individual, and attributes of the individual being trusted can all influence the nature of final trust assessments. While one individual ultimately makes his or her assessment of the trustworthiness of another, his or her decision cannot be understood in isolation from either the underlying social context and processes or the identities and attributes of the individuals involved. In the sections that follow we address these factors, highlighting their implications for the likelihood of trust failure.

Social Context Implications for Trustworthiness Assessments

While trust is touted as an efficient remedy for coping with social complexities and uncertainties (Luhmann, 1979), the social contexts within which trust tends to emerge render it, in at least three important respects, problematic. First, settings where trust is most needed—situations of interdependence, potential conflict, and risk (Thompson, 1967; Mayer, et al., 1995)—are those where betrayals of trust are known to be more prevalent. Research findings show that trust betrayals are more characteristic of individuals from certain occupational groups than others (Jones & Burdette, 1994). Self-reported propensities to betray trust relationships are greater than the norm for individuals in occupations with greater interpersonal contact and conflict—occupations such as school administrators, psychotherapists, sales persons, and attorneys. Betrayal propensities are lower than the norm for carpenters and computer operators. Interestingly enough, knowledge of the attendant risks of trusting in interdependent relationships is not sufficient to make people more cautious in vesting trust in others.

Second, contextual factors giving rise to the need for trust—physical distance, social distance, and asymmetric information/expertise (Shapiro, 1990, 1987; Zucker, 1986)—also make betrayals of trust much more feasible (Shapiro, 1987, 1990). Given asymmetric information and expertise (as frequently exists in attorney-client relations, accountant-CEO relations, machinist-manager relations, cross-functional team relations, and such), trusting individuals do not even have the skills and abilities needed to properly evaluate the performance of those being trusted. Given the inherent difficulties of monitoring, bonding and credentialling arrangements emerge as proxies for trustworthiness (Doctors are used that have proper credentials and malpractice insurance, and such). Still, these defenses pro-
vide limited protection against systematic abuses of trust—trusting individuals still have limited ability to tell whether those they trust are behaving responsibly and ethically. The sophisticated accountant can hide behind the protection of her or his professional designation (CPA) and intimate knowledge of accounting systems, using professional autonomy and freedom to access to embezzle funds (Cressy, 1953). Itinerant lawyers can manipulate billing records in ways that few but the most sophisticated of clients can detect (Lerman, 1990).

Third, accumulated trust appears to disarm otherwise vigilant principals, rendering such individuals increasingly helpless. While increased trust does provide a basis for reduced controls and vigilance (Ouchi, 1979), it does increase the risk that critical signs of untrustworthiness may simply be unobserved or ignored. Unfortunately, for contexts of overt subterfuge and deception, it is the prior accumulation of trust that affords sufficient latitude for deception to occur and go undetected (Blum, 1972). For deceivers, so long as conditions of trust are maintained (evidence to the contrary is hidden), trust can be easily manipulated for advantage. For situations of executive incompetence and unreliability—as occurs in situations of emergent alcoholism, substance abuse, and mental disorder—accumulated trust and latitude provide the conditions in which problem situations involving unreliability can remain undetected for extended periods of time.

For all practical intents and purposes, trust is a risky proposition—betrayal exists where trust is needed most, the factors giving rise to the need for trust facilitate betrayal, and accumulated trust serves to increase vulnerability. Beyond these factors, the social processes underlying the emergence of trust merit careful consideration.

**Process Considerations: Nature of the Trusting Process.**

Several features of the trust-building process open up real possibilities of trust failure—misattribution and betrayal. Difficulties stem primarily from the attribution and justification processes by which affect-based trust emerges and is maintained. First, the attribution processes governing trust decisions result in separation of trustworthiness assessments from the experiential base upon which they are formulated. Second, rationalizing and justifying processes allow individuals to minimize the salience of relationship partner shortcomings. Indeed, occasional evidence of untrustworthiness may paradoxically bolster trustworthiness assessments by forcing rationalizing individuals to commit greater cognitive resources to defending established beliefs and understandings.

**Decoupling Trustworthiness Assessments From Experiential Foundations**

As trust matures, we have argued, cognition-based trust becomes overlaid with affect-based trust and emphasis shifts from beliefs about partner predictability and dependability to beliefs about the stable personal qualities of relationship partners.
and their orientation to the relationship. This increase in affect-based trust brings needed reassurance to interdependent relationship partners. The process is marked by a gradual decoupling of trustworthiness assessments from the experiential foundations on which trust is initially based. Beliefs about the benevolence of relationship partners become taken as given and treated as no longer open to question (Rempel, et al., 1985). Unfortunately, this decoupling also opens up significant possibilities for deception and betrayal.

To some extent, the inherent affective nature of affect-based trust may provide explanation for why this decoupling takes place. As Zajonc observes, “once formed, an evaluation is not easily revoked…. Affect often persists after a complete invalidation of its original cognitive basis” (1980, p. 157). On a practical level, this decoupling of trust from behavioral foundations serves to make trust more resilient, less susceptible to the ebbs and flows of ongoing social interaction. This decoupling is also extremely efficient—once a relationship partner is seen as trustworthy and driven by sentiments of reciprocated care and concern, benevolent interpretations for all relevant behavior of the partner are possible—uncharitable interpretations need not even be considered. After all, according to this reasoning, good people do not do bad things, at least not intentionally and certainly not because of their nature (Holmes & Levinger, 1994). The affective nature of trust significantly structures and directs subsequent information processing in ways that favor benevolent interpretations of the behavior of trusted individuals. Given a-priori assumptions concerning the motives driving relationship partner behavior, considerable poetic license may be taken articulating retrospective accounts that make sense of partner behavior (Murray & Holmes, 1994).

This decoupling of trustworthiness assessments from experiential foundations is facilitated by behavioral changes that accompany adoption of the norms governing affect-based trust relationships. Where affect-based trust prevails, relationship partners adopt a need-based approach to responding to one another. Partner difficulties are responded to with added assistance, and the assistance provided is understood to be tangible evidence of the social bond in place. Whereas, under conditions of cognition-based trust alone, partner shortcomings might have stimulated questioning of partner trustworthiness, under conditions of affect-based trust, no such reflection and assessing necessarily occurs. Instead, partner unreliability, combined with “appropriate” responding, provides the foundation for a caring dependence relationship that hides real untrustworthiness. Especially where excuses are available to discount and minimize the importance of partner shortcomings, as discussed below, affect-based trust appears particularly susceptible to failure.

**Social Justification Processes that Perpetuate Failed Trust**

The gradual process by which trustworthiness assessments are decoupled from their foundations in experience does not render behavior irrelevant. While trustworthiness assessments are less influenced by immediate changes in peer conduct,
overall behavioral consistency with the assumptions underlying affect-based trust is still important. Given the level of effort needed to build trust relationships, the level of ego-investment in affect-based trust relationships that takes place (Kelley, 1979), and the substantial costs associated with distrusting, rationalizations and justifications for observed behavior are extremely important.

From recent research on trust in close relationships we are learning much about the attribution processes that function to preserve trust relationships (Holmes, 1991). Behaviors of relationship partners inconsistent with need-based responsiveness assumptions are disruptions that must be addressed. Core beliefs about relationships are protected by attribution processes that discount negative behavior and relegate it to lower levels of concern. Discounting processes take a number of forms including rejection, reconstrual, and refutation (Murray & Holmes, 1994).

Rejection. Denial of the significance of inappropriate behavior is possible when the behavior can be explained as a temporary perturbation brought about by situational factors. Partner unfriendliness can be understood as the product of unusual business and stress at work (Holmes & Levinger, 1994). Denial is also possible in contexts of considerable ambiguity where preferred interpretations cannot be immediately contradicted. For example, Fawn Hall, Oliver North's secretary during the Iran-Contra crisis, was able to deny to herself that her boss was doing anything wrong and justified her cooperation in the altering and shredding of top-secret documents—"He was a man whose character I knew well. I knew there was a reason why he was doing this." ("Fawn Hall," 1989). Where desired, trust can provide ample reason for denying the wrongdoing of a relationship partner.

Reconstrual. Reconstrual of apparent inappropriate behavior in a positive light—seeing virtues in faults—represents a second approach to coping with apparent undesirable relationship-specific behavior (Murray & Holmes, 1993). Stubbornness, suggestive of unresponsiveness in a relationship partner, may be reinterpreted as a sign of conviction and integrity. A peer's inability to complete assigned work during working hours can be interpreted as a sign of commitment rather than incompetence. Similarly, an employee's willingness to forego vacations may be interpreted as commitment rather than a desire to manage correspondence and cover-up embezzlement.

Refutation. When faults cannot be denied or recast in a positive light, efforts are made to interpret them within a larger context so that their specific effects on assessments are minimized. Partner faults can be acknowledged but coupled with insight into other important strengths with "yes, but..." arguments. Murray and Holmes (1993) found individuals high in trust likely to integrate their thoughts about negative aspects of a relationship with thoughts about positive qualities. Holmes and Rempel (1989) found highly trusting individuals to evaluate partner behavior and motives more positively after being asked to recall negative informa-
 tion about the partner than when asked to recall positive information. Murray and Holmes (1996) postulate that positive representations (beliefs and illusions rather than true perceptions) arise, not in spite of partner faults but precisely because of them.

These three sensemaking approaches—rejection, reconstrual, and refutation—yield rationalizations for discounting apparent evidence of untrustworthiness. They function in harmony with one another, allowing reason for maintaining trust beliefs in the face of considerable evidence to the contrary. As a practical example of the convergence of these approaches, consider the case of Congresswoman Enid Greene Waldholtz whose husband successfully forged checks, embezzled campaign funds, and channeled funds illegally to her election campaign. The case involved millions of dollars and extended over a five year period. Mrs. Waldholtz is the first to admit that “trust” was the culprit—it was because she trusted too much, and believed too strongly that he cared for her, that she was so easily deceived (“News Conference to discuss..,” 1995). When initially approached by staffers with evidence of Joseph Waldholtz’s misbehavior, she was able to deny the accusations with explanations provided by her husband. She describes Joe as a “seamless liar with an uncanny knack for talking people into doing what they wouldn’t normally do” (“Tearful Waldholtz..,” 1995). His explanations for irregularities in accounts were certainly enough for her. She easily discounted “mistakes” made by Joe to overwork—“I thought the problems were simply a matter of Joe being overextended…. I thought that Joe’s errors on the FEC report were due to the strain of trying to deal with his mother’s difficulties” (“News Conference to discuss..,” 1995). Joe’s efforts to restrict Enid’s access to information that would clue her into what was really happening (receiving all phone messages and screening calls, organizing her desk, handling all mail and invoices, and such), were interpreted as indicative of his true dedication and commitment. This is tantamount to seeing virtues in faults. While Enid was aware that others had reservations concerning Joe, she was able to consider his strengths in addition to his weaknesses. Indeed, she was inclined to tease her husband that “she would silence his critics by revealing his sweet private side” (“Tearful Waldholtz..,” 1995). Emphasizing compensating strengths appeared to bolster Enid’s confidence in Joe. Ultimately, each approach to excusing inappropriate behavior was used with effect, and considerable damage followed.

Murray, Holmes, and Griffin (1996) contend that there is much benefit in the maintenance of positive illusions within close relationships. These researchers find that, in general, intimates in close relationships do tend to see their partners as they wish to see them, through the filters provided by their ideals and rosy self-images, and that this is particularly true of satisfied relationship members. These researchers argue that idealization and poetic license work together to protect relationships from the harsh realities of interpersonal discovery that assuredly are encountered as relationships progress from infancy to maturity.
Potential benefits of positive illusions notwithstanding, it seems imperative that likely negative consequences of self-deception through excuse-making be clearly acknowledged. First and foremost, as noted above, unwise trust decisions can be made and considerable loss for trusting individuals as well as innocent third-parties can follow. Further, excuse-framing processes can work side by side with commitment escalation processes (Brockner, 1992) to entrap trusting individuals within dysfunctional and abusive relationships (Strube, 1988). As commitment within affect-based trust relationships increases there is likely to be a concomitant narrowing of attention that takes place and blinds decision makers to changing conditions and alternative courses of action. Given incentives to defend the status-quo and the belief that viable alternatives do not exist, change is unlikely.

**Individual Differences I: Attributes of Trusting Individuals**

It is important to note that some individuals are more inclined to make excuses for the behavior of relationship partners. There is considerable interest among management scholars in the role that individual differences in trust play in the development of interpersonal trust relationships (Mayer, et al., 1995). Interestingly enough, foundational research on dispositional trust suggests that trusting individuals are no more gullible and susceptible to betrayal than distrusting individuals (Rotter, 1980). In general, social context factors such as successful interaction in the past, rather than individual differences in generalized trust of others, are found to better explain interpersonal trust directed at specific individuals (Stack, 1988; Tardy, 1988). Other individual differences, however, such as levels of self esteem (Brockner & Rubin, 1985) and interpersonal attachment orientation (Hazan & Shaver, 1987) have been shown to be associated with inappropriate trusting and the propensity to rationalize away evidence of untrustworthiness (Strube, 1988).

Research on commitment escalation has shown that individuals with low self-esteem may be more inclined to rationalize and make excuses for others (Brockner & Rubin, 1985; Strube, 1988). For individuals with low self esteem, alternatives to rationalizing may appear particularly risky, the likelihood of alternative courses of action working out may appear low, and the true costs associated with continued investment in the relationship may be understated (Strube, 1988).

In a similar manner, recent research on adult attachment styles and their implications for behavior indicates that individuals with anxious/ambivalent proclivities (as distinct from secure or avoidant tendencies) may be more inclined to develop unhealthy trust relationships with others (Hazan & Shaver, 1990; Shaver & Hazan, 1994; Feeny & Noller, 1990; Hardy & Barkham, 1994). Drawing upon work from developmental psychology on the nature of parent-child relations, Hazan and Shaver (1987) posit that at least three distinct interpersonal orientations characterize adult attachment relationships: secure, where individuals find it relatively easy to get close to others, are comfortable depending on them, and do not
have concerns about being abandoned or about a partner’s getting too close; 
**avoidant**, where individuals are somewhat uncomfortable being close to others, find it difficult to trust and depend on them, and accordingly maintain healthy distances; and **anxious-ambivalent**, where individuals feel that others are reluctant to get as close as would be desired, feel neglected by partners and fear abandonment by partners, and recognize that the desire for closeness sometimes scares people away. These three hypothesized attachment orientations have been found to characterize adults as well as children (Hazan & Shaver, 1987), to distinguish among approaches to work and interpersonal relations in organizations (Hazan & Shaver, 1990), and ultimately to influence tangible work outcomes such as income (Hazan & Shaver, 1990; Hutt, 1992).

Individuals with anxious-ambivalent orientations represent a unique set of individuals prone towards commitment escalation and inclined towards excuse-making. Hazan and Shaver (1990) found anxious-ambivalent individuals to be preoccupied with attachment concerns, to expect coworkers to undervalue them (indicative of low self-esteem), to be motivated by the approval of others, and to worry that others would not be impressed with their work or would reject them. Further, anxious-ambivalent individuals expressed a greater desire to work with others and were slightly more inclined to report romantic interest in coworkers. Feeney and Noller (1990) show that, as attachment relationships become closer, anxious-ambivalent individuals become more distinct from secure and avoidant individuals, expressing greater preoccupation with, as well as emotional dependence and reliance upon, relationship partners. In-toto, strong incentives exist for anxious-ambivalent individuals to perpetuate trusting relationships and to defend trustworthiness assumptions when others would abandon such assumptions.

In general, research evidence indicates that approximately 20 percent of the population consists of anxious-ambivalent individuals, 55 percent secure, and 25 percent avoidant. However, where interdependent working relationships are concerned, the percentage of anxious-ambivalent individuals is likely to be greater avoidant individuals frequently self-select work roles with limited interdependence and anxious-ambivalent individuals seek out such roles (Hazan & Shaver, 1990). Of present concern is the fact that anxious-ambivalent individuals represent easy prey for would-be deceivers.

**Individual Differences II: Attributes of Trusted Individuals**

As a final consideration in our trust failures framework, it should be noted that attributes of an individual being trusted can influence the nature of trust relationships and the potential for betrayal. This is probably the area that has been researched least, but where great room for theoretical and empirical contribution exists. To this point we have been inclined to believe that it is extremely difficult to maintain a deception in close personal relationships (Luhmann, 1979). Yet,
given the fact that only a small portion of the crime that takes place in organizations actually is detected (Gabor, 1994) and our understanding of trust’s role in white-collar crime (Shapiro, 1990), it appears that successful deception does occur, and that some people may be better at it than others.

Hogan and Hogan (1994), in a summary report drawing on psychometric data and many interviews, identify several prototypical characteristics of the “ideal betrayer.” First and foremost, this person is charming, attractive, interesting, and well skilled in the art of flattery and ingratiation—the sort of individual inclined towards the development of more personal relationships. Further, this individual compulsively seeks the approval of others—the sort of individual inclined towards acting in ways that win others’ acceptance, perhaps through performing acts that address others’ needs. Finally, this individual has an unusual degree of self-absorption, egocentrism, or selfishness.

The profile of the “ideal betrayer” proffered by Hogan and Hogan’s analysis suggests an individual who selectively builds and nurtures what appear to be close relationships of trust. From the partner’s perspective, the relationship being developed could have definite qualities of apparent interpersonal care and concern. Individuals unwittingly in the favor of an “ideal betrayer” would receive help and assistance from the individual that others might not receive. The challenge might be to make sense of the individual’s inconsistency in behavior across relationships. Nevertheless, as logical explanations for these exceptions are maintained, we might expect a strong sense of trust to persist, and for subsequent signs of personal betrayal to go unnoticed. Ultimately, the betrayer’s self absorption allows him or her to manipulate relationships without careful consideration of the long term impacts of betrayal on relationship partners.

Summary and Implications of the Trust Failures Framework

Our analysis highlights the potential for trust failures. Trust relationships need not break down in order for trust failures to take place. Indeed trust failures take place whenever trust relationships are taken advantage of as well as when trust is inappropriately vested by trusting parties in relationship parties, and these relationships can persist as long as trusting parties are unaware of, or unwilling to admit to, the failings of trust relationships. Ambiguities in the context of trust make breakdowns difficult to detect and attribution processes underlying trust’s emergence augur against the detection of trust breakdowns. What we know about individual differences suggests that individuals with low self esteem and an anxious-ambivalent attachment orientation are more inclined to place themselves within situations that involve trust, more inclined to prematurely place their confidences in others and more inclined to use sense-making tactics to protect trust relationships in question. Finally, we now recognize that some individuals are more capable of maintaining deceptions than others, and that it may be difficult for dependent individuals to separate out
savvy impression management and the appearance of faithfulness from real trustworthiness. Important implications for research as well as practice follow from this analysis.

Perils of Unabated Enthusiasm for Trust

First and foremost, in contrast to present enthusiasm among management scholars for exploring the virtues of trust, it is important to recognize that unabated enthusiasm for trust poses risks. Barber (1983), Deutsch (1960), and Rotter (1980) all observe that there are times when distrust is very much appropriate and trusting would be foolhardy. Barber argues that the effectual nature of trust is extremely weakened if those who become justifiably distrustful do not have recourse to the law and other controls. Within organizations especially, monitoring, auditing, and insurance arrangements can function to forestall and compensate for failures in fiduciary responsibilities and are necessary complements to trust. Similarly, Luhmann (1979) argues that “Above all...trust depends on the inclination towards risk being kept under control and on the quota of disappointments not being too large.” Importantly, this can only be possible where appropriate sanctions are in place to address the possibilities of opportunistic conduct and betrayal. Future research must address the important role that constraints and protections have in emergent and continuing trust relationships.

The Satisfactions of Somewhat Less than Complete Interpersonal Trust

Second, the present analysis points to our lack of understanding of relationships of interpersonal confidence that are intermediary in nature, somehow characterized by something less than full trust and something more than complete distrust. Our general understanding is that trust relations tend towards stable reciprocated trust or stable reciprocated distrust, and that intermediary conditions are unstable in nature (Lewis & Wiegert, 1985; Fox, 1984). Understandably, as Barber (1982) argues, “perhaps because of its very importance we tend to exaggerate both the need for full trust and the evils of imperfections in trust processes” (1983, p. 166). Nevertheless, Barber admits that most of the time we get along well with less than complete trust.

In order to understand the nature of objective reality as distinct from idealized states, it is important that we extend our understanding of social relations characterized by intermediary levels of trust and come to recognize the potential for stability within such social relations. There are practical examples of stable admixtures of trust and distrust in impersonal relations that merit closer examination. For instance, Dawes and Thaler (1988) give the example of rural areas around Ithaca where farmers market fresh produce by means of unattended produce stands by the side of the road. Dawes and Thaler observe, “There is a cash box on
the table and customers are expected to put money in the box in return for the vegetables they take. The box has just a small slit, so money can only be put in, not taken out. Also, the box is attached to the table, so no one can (easily) make off with the money" (1988, p. 45). This example suggests a willingness to trust as well as a seasoned understanding of human nature.

Moving beyond impersonal relations among farmers and customers, is it possible for stable personal relations of confidence to be characterized by something less than complete trust and something more than complete distrust, for distrust and trust to coexist, or for distrust to enter into trusting relationships without “tainting” the relationship? In parent-child relations it seems reasonable for children to trust their parents in most matters while parents may be more discerning, trusting their children in certain matters but knowing better than to trust them in other matters (Baier, 1985; Held, 1990). Likewise, in superior-subordinate relations, the potential for superiors to have constrained trust in subordinates that lack experience, practical knowledge, and perspective seems possible and appropriate.

In everyday life we recognize the value of realism in interpersonal relations. We are inclined to respect mature trust relationships in which knowledge of trust’s limitations does not place the complete relationship in peril. Nevertheless, the foundations of current research on trust relationships in organizations do not provide for systematic consideration of this possibility. Further research addressing the resilience of trust relations, the conditions under which intermediary conditions of trust are possible, and the conditions under which trust and distrust can coexist, is definitely needed.

Not all Forms of Trust are Created Equal

One important insight from the present analysis concerns the differential susceptibility of alternate forms of interpersonal trust to betrayal and abuse. Affect-based trust appears particularly susceptible, given (1) its emphasis on attributions regarding the motives of relationship partners and the potential for trustworthiness assessments to become decoupled from the experiential foundations upon which they are based, and (2) the substantial emotional investments it entails and the social justification processes that can perpetuate it.

This insight into the differential susceptibility to alternate trust forms to betrayal and abuse extends our understanding of the differences between cognition- and affect-based trust. Increasingly, we are understanding that a singular emphasis on level of trust, apart from consideration of the substantive nature of trust, is inappropriate (McAllister, 1995). Prior research indicates that the antecedents and consequences of distinct forms of interpersonal trust differ in important ways (Lewicki & Bunker, 1995; McAllister, 1995). Insights from the present conceptual analysis extend this understanding and merit further empirical attention.
Individual Differences and Trust

A rich tradition of research on interpersonal trust as an individual difference, which was precipitated by the foundational work of Rotter (1967, 1971), exists. Management scholars today recognize individual differences in trust as a likely antecedent of trust within specific relationships (Mayer, et al., 1985). Increasingly, however, we understand that individual differences in propensities to trust are best suited to explaining behavior in novel situations, and that the outcomes of interaction over time are better predictors of trust in stable social situations and ongoing working relationships (Stack, 1988). While interpersonal trust as an individual difference may be of limited significance, other individual differences may have profound implications for trust in ongoing working relationships. In particular, adult attachment styles (Hazan & Shaver, 1990; Hardy & Barkham, 1994) appear to influence the dynamics of ongoing trust relationships in important ways.

Research evidence reviewed here indicates that individuals with anxious-ambivalent attachment styles, as distinct from individuals with secure or avoidant styles, appear more inclined to seek out interdependent working relationships, to be preoccupied with establishing trusting relationships, to prematurely express trust and confidence in relationship partners, and to defend trust relationships in the presence of evidence bringing trustworthiness assumptions into question. Further research addressing the factors influencing adult attachment styles is needed. While developmental psychologists argue that infant parent attachment styles persist into adulthood, it seems likely that attachment experiences of significance in later life, including attachments of individuals to work groups and organizations, may have implications for the attachment orientation that characterizes subsequent interpersonal attachments. That is, recent breakdowns in the psychological contracts connecting individuals with organizations may have the effect of increasing the salience of interpersonal attachments and possibly transforming secure orientations into anxious-ambivalent orientations (Morrison & Robinson, 1997; Robinson, 1996; O'Neill & Lenn, 1995). Further research addressing these and other factors influencing adult interpersonal attachment orientations may contribute much to our understanding of the role of individual differences in trust relationships.

CONCLUSION

Given the potential dysfunctions of trust relationships, the work of managing trust relationships and facilitating the emergence of functional trust and distrust relationships emerges as a new managerial challenge, as well as an important area of scholarly inquiry. Baier (1987), describing parent-child relations, observes that loving parents create a nurturing environment that supports the emergence of trust, but prepares children to distrust on appropriate occasions as well as trust. Interest-
The Second Face of Trust

ingly enough, done effectively, parenting serves to promote the emergence of functional confidence relations and minimize the negative consequences of distrust for social solidarity (discretion in the application of distrust only to situations where betrayal is both possible and probable is learned). Could it be that the capacity for effectively distrusting is a necessary precondition for the development of effective trusting relationships? Based upon what we have known for some time about the efficiencies of trust, as well as what we now know about the darker side of trust, we conclude that leaders contribute much by preparing followers to be effective distrusters specifically because effective trust relationships are so important.

NOTES

1. It is instructive to note that the emphasis of Murray, Holmes, and Griffin (1996) on the functional aspects of positive illusions is in one sense very characteristic of the excuse making strategy offered by Homes and Levinger (1994). Dysfunctions of illusions are denied, what appears to be a fault is now interpreted as a strength, and the ultimate result of exploring the benefits of positive illusions with sufficient insistence is that reason is found for preferring positive illusion over some more objective reality.

REFERENCES


News Conference to discuss Marital and Professional Problems. (1995, December 11) FDCH Political Transcripts, Salt Lake City, Utah.


The Second Face of Trust


