INITIAL TRUST FORMATION IN NEW ORGANIZATIONAL RELATIONSHIPS

D. HARRISON McKNIGHT
Florida State University

LARRY L. CUMMINGS
NORMAN L. CHERVANY
University of Minnesota

Arguably, the most critical time frame for organizational participants to develop trust is at the beginning of their relationship. Using primarily a cognitive approach, we address factors and processes that enable two organizational parties to form relatively high trust initially. We propose a model of specific relationships among several trust-related constructs and two cognitive processes. The model helps explain the paradoxical finding of high initial trust levels in new organizational relationships.

Several trust theorists have stated that trust develops gradually over time (e.g., Blau, 1964; Rempel, Holmes, & Zanna, 1985; Zand, 1972), but when contrasted with some recent empirical findings, their theories present an interesting paradox. By positing that trust grows over time, these trust theorists implicitly assume that trust levels start small and gradually increase. Some researchers, then, expecting this, have been surprised at how high their subjects' early trust levels were—both in survey and experimental studies (e.g., Berg, Dickhaut, & McCabe, 1995; Kramer, 1994). For example, economics-based researchers Berg et al. (1995) expected subjects to exhibit low to medium trust in each other when faced with a trust dilemma. Instead, their subjects frequently exhibited high-trust—passing to a second subject dollars they were given during the first part of the experiment, without any reason to expect their generosity to be reciprocated. Another example is Kramer's (1994) survey of MBAs who were previously unknown to each other. Because the MBAs had no interaction history, one would have expected them to have low trust levels; however, surprisingly, Kramer found high trust levels among these individuals.

THEORETICAL BACKGROUND

The Paradox of High Initial Trust Levels

High initial trust findings are paradoxical because, as we stated above, several trust theorists predict low initial trust. By "initial" we mean when parties first meet or interact. An example of initial trust predictions is provided by economics- or calculative-based trust researchers (e.g., Coleman, 1990; Williamson, 1993), who theorize that individuals make trust choices based on rationally derived costs and benefits (Lewicki & Bunker, 1995; Shapiro, Shepard, & Cheraskin, 1992). Thus, calculative-based trust theorists would predict that the lack of incentives (benefits) of Berg et al.'s subjects would result in low levels of trusting behavior among them. However, Berg et al.'s results do not agree with this prediction. As another example, knowledge-based trust theorists propose that trust develops over time as one accumulates trust-relevant knowledge through experience with the other person (Holmes, 1991; Lewicki & Bunker, 1995). From this perspective, Kramer's study participants would require time and an interaction history to develop a high level of trust in each other. However, Kramer's results directly contradict this. Thus, studying initial trust formation is important, because the results from such studies require an explanation beyond what calculative-based and knowledge-based trust theories provide.

In this article we argue that the paradox of high trust in initial relationships may be ex-
plained by identifying "hidden" factors and processes that enable trust to be high when people in organizations first meet. We develop a model of initial trust formation to explain why trust may be high when members of organizations barely know each other.

Increasingly Common New Work Relationships

In today's work environment, interacting with a new manager or with new coworkers is becoming commonplace. Such a situation involves initial trust, because the parties have not worked together long enough to develop an interaction history. Initial trust situations occur naturally, when an employee or manager is newly hired or transferred, when cross-functional teams are formed, when salespeople or consultants call, when mergers bring two sets of employees together, or when a new joint venture begins.

These situations are becoming more common because of increased mergers or acquisitions and because widespread corporate restructuring and reduced employee loyalty have increased the typical turnover rate of organizational workers and managers (e.g., Evans, Gunz, & Jalland, 1996). The nature of tasks is also increasing new work encounters, as temporary task teams or project engagements become the norm in organizations. Meyerson, Weick, & Kramer (1996) explain that such environmental factors as outsourcing and labor/skill shortages have increased the number of temporary task teams. Further, communication technology now enables millions of people (Henry & Hartzler, 1997) to work on virtual task teams (Lipnack & Stamps, 1997), in which participants are often new to each other.

Because working together well requires some level of trust (Bromiley & Cummings, 1995), increasingly common new work encounters demand that the parties come to trust each other quickly (Meyerson et al., 1996). Thus, the need exists for a model of how trust initially forms.

A Model of Initial Trust Formation

Initial trust between parties will not be based on any kind of experience with, or firsthand knowledge of, the other party. Rather, it will be based on an individual's disposition to trust or on institutional cues that enable one person to trust another without firsthand knowledge. Figure 1 depicts the initial trust formation model. The model applies only to new encounters between people. Therefore, it excludes experiential processes (e.g., observing the trustees' behavior), but it does include cognitive processes and factors that lead to initial trust. As we explain in the next section, the model uses constructs from four trust research streams.

We define "trust" to mean that one believes in, and is willing to depend on, another party (e.g., Mayer, Davis, & Schoorman, 1995). This high-level trust concept can be broken into two constructs: (1) trusting intention, meaning that one is willing to depend on the other person in a given situation (e.g., Currall & Judge, 1995), and (2) trusting beliefs, meaning that one believes the other person is benevolent, competent, honest, or predictable in a situation (Mayer et al., 1995).

Distinguishing two constructs that constitute trust is important, because the word "trust" is so confusing (Shapiro, 1987a) and broad (Williamson, 1993) that it almost defies careful definition (e.g., Gambetta, 1988). The distinction between trusting beliefs and trusting intention follows the Fishbein & Ajzen (1975) typology separating constructs into beliefs, attitudes, intentions, and behaviors. We exclude attitudes and behaviors here so we can focus the article on cognitive concerns.

At a summary level, the model (Figure 1) implies that trust forms because of one's disposition to trust, one's institution-based trust, and two cognitive processes we discuss later. "Disposition to trust" refers to a tendency to be willing to depend on others. "Institution-based trust" means that one believes impersonal structures support one's likelihood for success in a given situation. To avoid the vagueness of discussing the formation of the broad trust concept, we use the model to explore the specific formation of trusting intention and trusting beliefs.

Theoretical Foundations for the Initial Trust Model

A significant body of knowledge from five research streams sheds light on how initial trust forms. In addition to knowledge-based and calculative-based trust research, three other research streams have been used: (1) personality based, (2) institution based, and (3) cognition
based. According to personality-based trust researchers, trust develops during childhood as an infant seeks and receives help from his or her benevolent caregiver (Bowlby, 1982; Erikson, 1968), resulting in a general tendency to trust others (Rotter, 1967). Institution-based trust researchers maintain that trust reflects the security one feels about a situation because of guarantees, safety nets, or other structures (Shapiro, 1987a; Zucker, 1986). Cognition-based trust researchers purport that trust relies on rapid, cognitive cues or first impressions, as opposed to personal interactions (Brewer, 1981; Lewis & Weigert, 1985b; Meyerson et al., 1996).

The personality-, institution-, and cognition-based research streams each can help explain the paradox of high initial trust levels. Personality-based trust theorists, for example, would say that subjects in Kramer's study with a high disposition to trust would have high initial trust. However, this may not explain the overall high level of trust, because it is unlikely that nearly all of Kramer's subjects had high disposition-to-trust levels. Institution-based theorists would argue that the structure of the classroom situation supported high levels of initial trust by enabling subjects to feel secure in the situation (e.g., Lewis & Weigert, 1985a; Shapiro, 1987a). Kramer himself used a cognitive explanation for his results, attributing high trust levels to the favorable views the MBAs had toward those of their own kind.

Each of these three trust research streams can partially explain high initial trust, but if we focus on only one of the three, we face the danger of the other two acting as hidden confounds. The danger of hidden confounds exists because, in a given context, all three types of factors—personality, institutional, and cognitive—may be present. Sitkin and Pablo (1992) demonstrated this with respect to the prediction of risk behavior; they identified a hidden personality-related construct (risk propensity) that explained the paradoxical empirical results researchers found when predicting risk behavior with situational and organizational variables. By identifying the hidden construct, Sitkin and Pablo were able to specify more fully the antecedents of risk behavior.

To explain the paradox of high initial trust, we will justify the initial trust formation model (Figure 1), based on the detailed constructs and pro-
cesses in Figure 2. The detailed model's constructs and processes come from four of the five research streams mentioned above (see Figure 2): (1) personality (faith in humanity), (2) institutional (institution-based trust), (3) calculative (trusting stance), and (4) cognitive (categorization processes and illusions of control process). Trusting beliefs and trusting intention come from more than one research stream. The fifth research stream—knowledge-based trust—assumes that the parties have firsthand knowledge of each other, based on an interaction history. Because "initial relationships," by definition, have no interaction history, firsthand knowledge-based trust does not apply to them. Hence, firsthand knowledge-based trust formation processes lie outside the scope of this article. We address second-hand knowledge, such as reputation, as a categorization process.
We chose the model’s four trusting beliefs because they are the most commonly used trusting beliefs in the literature (e.g., Mayer et al., 1995). Each trusting belief is proposed to be more highly correlated with other trusting beliefs than with the model’s other constructs, indicating a type of convergent validity. The constructs in institution-based trust and disposition to trust also display convergent validity. Trusting beliefs are proposed to be more highly related to trusting intention than they are to institution-based trust and disposition to trust, as reflected by the box grouping them together as a composite trust concept.

We define the model’s constructs at the individual level of analysis. Hence, the model’s constructs are internally consistent, even though one set of constructs (institution-based trust) reflects group- or organization-level phenomena. In limiting the article’s scope to the individual level of analysis, we are not disallowing social and organizational effects. In fact, we believe organizational-level constructs affect trust. However, these lie outside the scope of the article.

Our primary contribution is in creating a model that explains why trust can be high initially. But the model also addresses a second paradox. Although trust frequently has been termed “fragile” (e.g., Worchel, 1979), it has also been described as “robust” (e.g., Luhmann, 1979). An added contribution is that we discuss the model’s implications for understanding why trust is considered both fragile and robust. We include research implications for the model in the last section of the article.

A MODEL EXPLAINING THE HIGH INITIAL TRUST PARADOX

Disposition to Trust Affects Trusting Beliefs and Trusting Intention

A person exhibits a disposition to trust to the extent that she or he demonstrates a consistent tendency to be willing to depend on others across a broad spectrum of situations and persons—a personality construct that builds on the work of Erikson (1968) and Rotter (1967, 1971, 1980). In contrast to others, we distinguish between two types of disposition to trust, each of which affects trusting intention in a different way: (1) faith in humanity and (2) trusting stance. Reflecting the traditional view of personality-based trust, “faith in humanity” means that one believes that others are typically well-meaning and reliable (e.g., Rosenberg, 1957; Wrightsman, 1991). “Trusting stance” means that one believes that, regardless of whether people are reliable or not, one will obtain better interpersonal outcomes by dealing with people as though they are well-meaning and reliable. Because it reflects a conscious choice, trusting stance derives from the calculative-based trust research stream (e.g., Riker, 1971). As an example of trusting stance, one of the authors asked a respondent if she trusted her new manager, who had just been hired from outside the company. She replied “yes,” explaining that she generally trusts new people until they give her some reason not to trust them. Hence, she exhibited a personal strategy to trust newcomers. Both faith in humanity and trusting stance are “dispositional,” which refers to personal tendencies, not traits, because they reflect tendencies that apply across various situations.

Mixed empirical findings. Researchers have experienced mixed results when using disposition to trust to predict trust. From their research, Johnson-George and Swap conclude that constructs like disposition to trust “do not accurately determine an individual’s trust in another under particular circumstances” (1982: 1007). In contrast, others have found disposition-related constructs to be important predictors. For example, Goldstein, Schorr, and Goldstein (1989) found that disposition-related trust had a statistically significant effect on peoples’ mistrust of federal nuclear plant authorities whom they did not know personally. Mayer et al. (1995) reviewed additional evidence that disposition-based trust is important to other trust constructs. However, Holmes (1991) declares that although researchers assume disposition to trust is a contributor to the development of relationship-specific trust, this link has not been proven.

Resolving the research findings. We can resolve these mixed results by using our initial trust model. The time frame of the relationship is important in predicting the effects of disposition to trust. Although other variables will swamp the effects of a person’s trusting tendency in ongoing relationships, disposition to trust likely will have a significant effect on a person’s trusting beliefs and trusting intention in new organizational relationships. Johnson-George and
Swap cite evidence that disposition to trust predicts what they call "trusting behavior" only when parties are new to each other in "highly ambiguous, novel, or unstructured situations, where one's generalized expectancy is all one can rely on" (Johnson-George & Swap, 1982: 1307; cf., Rotter, 1980). A transferred worker's relationship with a new manager in an unfamiliar area exemplifies such a situation, because roles and relationships are not yet clear. Goldstein et al.'s (1989) finding—that dispositional trust was related to specific trust of federal authorities individuals had never met—supports the idea that dispositional trust is salient when people do not know each other.

Faith in humanity effects. Because faith in humanity reflects the extent to which one believes that nonspecific others are trustworthy, this characteristic will probably affect one's initial trusting beliefs (Kramer, 1994). Already-developed patterns of thinking about relationships in general tend to transfer to a specific initial relationship. This is particularly true if an individual cannot draw on other reasons (e.g., trusting beliefs or institution-based trust), because the situation, the type of relationship, and the type of other person are new. In other words, if no more specific situational information is available, one will rely on one's basic beliefs about human nature (Wrightsman, 1991), as reflected in faith in humanity. This is similar to the argument, presented by Mullins and Cummings, that "weak situations" display ambiguity in terms of the meaning of the situation (in press). In weak situations the person's disposition will be more salient than the situation. Initial trust-related situations may be ambiguous because the parties' roles or task may be new. Rotter (1971) has said that the novelty of the situation affects how salient dispositional trust will be. In a novel situation, then, faith in humanity will enable trusting beliefs to be high.

Proposition 1: In the initial relationship, to the extent the situation is novel and ambiguous, faith in humanity will lead to trusting beliefs.

Trusting stance effects. Trusting stance influences one to be intentionally willing to depend on another, regardless of beliefs in the other. An individual with high trusting stance probably believes that things turn out best when one is willing to depend on others, even though others may or may not be trustworthy. Thus, trusting stance does not lead to beliefs about the other person; rather, it directly supports one's willingness to depend on that person.

Proposition 2: In the initial relationship, trusting stance will lead to trusting intention. The effects of trusting stance on trusting intention will not be mediated by trusting beliefs.

Institution-Based Trust Affects Trusting Intention

Institution-based trust means that one believes the necessary impersonal structures are in place to enable one to act in anticipation of a successful future endeavor (e.g., Shapiro, 1987a; Zucker, 1986). Trusting intention at the beginning of a relationship may be high because of high institution-based trust levels. In the literature two types of institution-based trust are discussed: (1) situational normality, defined as the belief that success is likely because the situation is normal, and (2) structural assurances, defined as the belief that success is likely because such contextual conditions as promises, contracts, regulations, and guarantees are in place. Later, we discuss the effects of institution-based trust on trusting beliefs. In this section we look at direct effects on trusting intention.

Situational normality belief effects. Situational normality belief stems from the appearance that things are normal (Garfinkel, 1963: 188) or "customary" (Baier, 1986: 245). or that "everything seems in proper order" (Lewis & Weigert, 1985a: 974). Situational normality involves a properly ordered setting that appears likely to facilitate a successful interaction. For example, a person who enters a bank tends to expect a setting conducive to both customer service and fiduciary responsibility that is reflected in the workers' professional appearance, the prosperous and secure physical setting, and the friendly, yet safe, money-handling procedures. The individual's belief that the situation is normal helps that person feel comfortable enough to rapidly form a trusting intention toward the other party in the situation.

Garfinkel's experiments, in contrast, demonstrated that when people face inexplicable, abnormal situations, trust between them breaks down. For example, one subject told the experi-
menter he had a flat tire on the way to work. The experimenter responded, “What do you mean, you had a flat tire?” (1963: 221). At this point trust broke down because of the experimenter’s illogical question.

Situational normality also can relate to one’s comfort with one’s own roles and the other person’s roles in that setting (Baier, 1986). Socially constructed roles create a shared understanding among members of the social system that facilitates trusting intention among them.

Proposition 3: In the initial relationship, situational normality belief will lead to trusting intention.

Structural assurance belief effects. Shapiro refers to structural safeguards in terms of institutional “side bets” (1987b: 204), such as regulations, guarantees, and legal recourse.

- Regulations enable people to feel assured about their expectations of the other party’s future behavior (e.g., Sitkin, 1995). For example, a company subcontracting the construction of metric-sized engine parts relies on the other party to use the same measures the company itself uses since they are specified by well-accepted metric standards makers.

- Guarantees mitigate the perceived risk involved in forming trusting intention. Hence, structural assurance belief leads to trusting intention, as evidenced by Zaheer, McEvily, and Perrone’s (in press) finding that structure-based trust is positively related to interpersonal trust.

- Legal recourse (i.e., regarding contracts or promises) is related to trusting intention for two reasons (Baier, 1986). First, the trustor feels comfortable that the promise has the type of significance in the particular setting so that the trusted person will make every effort to fulfill it, or risk reaping sanctions through social disapproval or legal action (e.g., Sitkin, 1995). Second, the trustor feels comfortable that the trusted person, out of either socially learned behavior patterns or fear of sanctions, will act according to the norms surrounding promise in the social setting.

Structural assurance belief will be more influential in the initial relationship than it will later, because information about the other person is very incomplete when the relationship begins, making situational information quite salient.

Proposition 4: In the initial relationship, structural assurance belief will lead to trusting intention.

Structural Assurance Belief and Situational Normality Belief Affect Trusting Beliefs

Structural assurance belief is likely to affect trusting beliefs for three reasons, the last two of which are shared by situational normality belief. First, believing that a situation is bounded by safeguards enables one to believe that the individuals in the situation are trustworthy. For example, a new employee can better believe in the boss’s benevolence if the employee believes the workplace has procedures that punish abusive managerial treatment. Also, an employee is more likely to believe a new coworker is competent if that individual believes the department’s hiring process is sound.

Second, the institutions in the situation reflect the actions of the people involved; therefore, beliefs about the institutions will help form beliefs about the people who are involved in the institutions. For example, within the bounds set by corporate practices and procedures, the boss is the chief administrator of fairness in the workplace. Thus, one’s belief about the structures supporting fairness in the workplace will support one’s belief in the boss’s benevolence. Similarly, a situational normality belief implies that the people in the situation will act normally and can therefore be trusted.

Third, based on many studies of cognitive consistency (Abelson et al., 1968), we believe that structural assurance and situational normality beliefs will probably stay consistent with related beliefs, such as trusting beliefs. Cognitive consistency is probably even more salient during the initial relationship, before beliefs about the other person and the situation become highly differentiated through experiential knowledge (Sitkin & Roth, 1993).

Proposition 5: In the initial relationship, the trusting beliefs will be a function of structural assurance belief and situational normality belief.

Trusting Beliefs Affect Trusting Intention

Evidence for the link between trusting beliefs and trusting intention is reviewed elsewhere (Mayer et al., 1995). Dobing (1993) also found a strong relationship between his trusting beliefs and trust (willingness to depend) constructs. Logically, if one believes that the other party is benevolent, competent, honest, and predictable,
one is likely to form a trusting intention toward that person. Therefore, trusting beliefs will positively impact trusting intention.

At a more general level, the literature that links beliefs to intentions supports this relationship (e.g., Davis, 1989; Fishbein & Ajzen, 1975). Ajzen (1988) discusses evidence that beliefs and intentions tend to stay consistent. They should be especially consistent at first, when one has no experiential basis not to believe the other person is trustworthy.

Proposition 6: In the initial relationship, trusting intention will be a function of benevolence belief, competence belief, honesty belief, and predictability belief.

Faith in Humanity and Trusting Stance Affect Structural Assurance Belief

Faith in humanity reflects one’s lifelong experiences with others (e.g., Rotter, 1967). A person who believes in the honesty and benevolence of people generally will probably have stronger beliefs in the security afforded by human institutions. Stated another way, one’s structural assurance belief is probably partly based on how one feels about people in general, because people play roles that relate to how secure the situation is. Hence, one’s feelings about people in general will likely influence one’s structural assurance belief. This is more likely to be true at the beginning of a relationship, when beliefs about the situation are based more on assumptions than on facts.

Trustung stance also will affect initial structural assurance belief. An individual with high trusting stance believes that trusting others facilitates success, regardless of his or her beliefs about specific people. This assumption is consistent with a perception that safeguards or safety nets will protect the individual involved from bad consequences other people cause. In other words, developing a high level of structural assurance belief is facilitated when one has a high level of trusting stance. When parties first meet, the person’s already-formed high level of trusting stance will tend to encourage a corresponding high level of structural assurance belief. Over time, the relationship between these constructs may be reciprocal, but at the initial meeting, a person will rely on his or her prior tendencies, in terms of trusting stance, to form structural assurance belief.

Proposition 7: In the initial relationship, faith in humanity and trusting stance will lead to structural assurance belief.

Categorization Processes Enable High Levels of Trusting Beliefs

In a new relationship a person may use three types of categorization processes to develop trusting beliefs: (1) unit grouping, (2) reputation categorization, and (3) stereotyping. “Unit grouping” means to put the other person in the same category as oneself. “Reputation categorization” means that one assigns attributes to another person based on second-hand information about the person. “Stereotyping” means to place another person into a general category of persons.

Unit grouping. Because those who are grouped together tend to share common goals and values, they tend to perceive each other in a positive light (Kramer, Brewer, & Hanna, 1996). Hence, one group member will be more likely to form trusting beliefs toward another group member. For example, Zucker, Darby, Brewer, and Peng (1996) found that being a member of the same organization generated the trust needed for scientists to collaborate on research. In their study Brewer and Silver (1978) found that people perceived in-group members to be more trustworthy than out-group members. These studies provide evidence that unit grouping quickly leads to high levels of trusting beliefs. As an example of the rapid effects of unit grouping, in studying prospective dating couples who had never met, Darley and Berscheid (1967) found that the knowledge that one would be paired with the other tended to enhance the former’s beliefs about the latter’s characteristics. Applied to a new task team, unit grouping would enable one member to immediately form trusting beliefs about another team member.

Reputation categorization. Those with good reputations are categorized as trustworthy individuals. Reputation may reflect professional competence (Barber, 1983; Powell, 1996) or the other trusting beliefs: benevolence (Dasgupta, 1988), honesty, and predictability. A person may be perceived as a competent individual because she or he is a member of a competent group
(Dasgupta, 1988) or because of her or his actions. Therefore, if the individual has a good reputation, one will quickly develop trusting beliefs about that individual, even without firsthand knowledge.

**Stereotyping.** Stereotyping may be done either on a broad level, such as gender (e.g., Orbell, Dawes, & Schwartz-Shea, 1994), or on a more specific level, such as prejudices for or against occupational groups (e.g., used-car salespeople). At their first meeting, parties may form stereotypes about each other, based on voice (e.g., male/female or domestic/foreign; Baldwin, 1992) or physical appearance (Dion, Berscheid, & Walster, 1972; Riker, 1971). By positive stereotyping, one can quickly form positive trusting beliefs about the other by generalizing from the favorable category into which the person was placed.

*Proposition 8: In the initial relationship, categorization processes that place the other person in a positive grouping will tend to produce high levels of trusting beliefs.*

**Illusions of Control Process: Interactive Effects That Elevate Trusting Beliefs**

Here we explain why the illusions of control process will interact with categorization processes, faith in humanity, and structural assurance belief to produce high levels of trusting beliefs. People in an uncertain situation will take small actions to try to assure themselves that things are under their personal control (Langer, 1975). This results in unrealistically inflated perceptions of personal control (Taylor & Brown, 1988), which Langer (1975) terms “illusions of control.” Illusions, obviously, involve perceptions that differ from reality, and considerable evidence demonstrates the presence of illusion in cognitive processing (e.g., Fiske & Taylor, 1984).

The illusions of control process that helps build trusting beliefs may be similar to the process by which people become overconfident of their judgments, as reported by Langer (1975) and Paese and Sniezek (1991). First, one forms a tentative belief, and then one watches for clues that confirm the belief. Even without evidence, the effort of watching tends to overinflate confidence in one’s judgment (Davis & Kotteman, 1994). Similarly, even a slight effort at confirming one’s tentative trusting beliefs in another may overinflate one’s confidence that high levels of trusting beliefs are warranted.

**Token control efforts.** People may make an initial effort to think about another person’s trustworthiness, or they may, upon meeting a person, immediately attempt to gauge whether or not they can influence that person in some small way (e.g., make them smile). We term such actions “token control efforts.” One individual is not trying to categorize the other but, rather, is testing his or her ability to deal with the other individual successfully.

Initially, a person is likely to use token control efforts because she or he does not know from experience whether or not the other has the attributes needed to be considered trustworthy. After making such small control efforts, the individual may form an unjustifiably strong confidence that one’s positive categorization, faith in humanity, and structural assurance beliefs are correct and, therefore, that the other person is trustworthy. Langer (1975) for example, in a study of overconfidence, found that token control efforts to improve one’s chances in a lottery (i.e., by choosing their own lottery ticket) made one very overconfident of winning.

Trust theorists have posited that trust building involves illusion (Holmes, 1991; Meyerson et al., 1996). The results of one empirical trust study support this position. Kramer (1994) found that ruminating for a few minutes about others’ motives and intentions increased a person’s confidence in the accuracy of his or her judgments of the others. Mental assessments tend to increase one’s confidence because, as Langer (1975) found, they move a task further from the realm of chance and closer to the realm of a skill-based judgment task. Hence, token control efforts can support a person’s confidence in trust-related beliefs.

**Interactive effects.** We propose that token control efforts will interact with categorization processes, faith in humanity, and structural assurance belief, strengthening an individual’s capacity to form trusting beliefs. Token control efforts will give a person the illusion that his or her positive faith in humanity can apply to the particular other party by convincing the individual that he or she is applying skill—not just chance—to a trust-related judgment of the other party. Similarly, token control efforts will
(1) build confidence that one’s positive categorization of the other person is correct, and (2) bolster confidence that structural safeguards make the environment secure and, by association, the people involved trustworthy.

**Proposition 9:** In the initial relationship, token control efforts will strengthen the tendency of categorization processes, faith in humanity, and structural assurance belief to produce high levels of trusting beliefs.

**MODEL IMPLICATIONS FOR TRUSTING INTENTION FRAGILITY/ROBUSTNESS**

Paradoxically, although trust often has been termed “fragile” (e.g., Worcel, 1979), it has also been described as “robust,” in that it progresses in upward spirals (e.g., Zand, 1972) or becomes more fully developed over time (Gabarro, 1978; Sitkin & Roth, 1993). Our model of initial trust development improves our understanding of how initial trusting intention, the model’s ultimate consequent, may be either fragile or robust under different conditions.

**Conditions Under Which Initial Trusting Intention Is Likely to Be Fragile**

Initial trusting intention is likely to be fragile under three conditions: (1) inadequate support from Figure 2’s antecedent constructs, (2) the tentative and assumption-based nature of the antecedent constructs, and (3) high perceived risk. Fragile refers to a trusting intention level that is likely to undergo large changes during a given time frame. That is, as time passes and conditions change, to what extent is the level of trusting intention likely to change? Although researchers’ use of the term “fragile” typically refers to a high trust level suddenly becoming low, we define the term so that it could apply to either high or low trust levels. At high levels, fragile trust is subject to rapid decreases, whereas at low levels, fragile trust is subject to rapid increases. In either case fragile trust is unstable, quickly changeable, or easily influenced.

Robust (the opposite of fragile) refers to a trusting intention level that does not change dramatically during a given time frame. As an example of the difference between robust and fragile, if a trusting intentions level were to move downward dramatically, we would call it fragile rather than robust.

**Inadequate antecedent support.** Intuitively, the more highly positive antecedents trusting intention has, the less fragile it will be; a weak combination of trusting intention antecedents will result in a fragile trusting intention. That is, if trusting intention is associated with only one highly positive antecedent, it is likely to change downward soon after the initial period of a relationship.

For example, suppose a home buyer with a low disposition to trust and low trusting beliefs in the home builder enters into a contract to have a house built, based on the belief that the legal processes will provide a safety net (structural assurance belief). Further, the home buyer has no cues by which to categorize the home builder positively. In this situation the only highly positive antecedent of trusting intention is structural assurance belief. In such cases, Dasgupta remarks that the trust one person has in another to fulfill a contract rests precariously upon the power of the agencies that are able to enforce contracts: “If your trust in the enforcement agency falters, you will not trust persons to fulfill their terms of an agreement and thus will not enter into that agreement. . . . It is this interconnectedness which makes such a contract a fragile commodity” (1988: 50). However, if the home buyer’s structural assurance belief is accompanied by trusting beliefs in the home builder, he or she may be able to maintain trusting intention even when his or her structural assurance belief wavers.

In general, then, when any particular antecedent of trusting intention is the only one at a high level, trusting intention most likely will be fragile. For example, the home buyer’s honesty belief is an antecedent of trusting intention that can crumble quickly (Dasgupta, 1988), if the buyer’s experience with the builder indicates that the belief is mistaken. We previously discussed how disposition to trust may be a weak trusting intention antecedent in the presence of a strong situation. The home buyer’s high disposition to trust will not hold up if he or she has reason to believe that structural assurances are missing or that the home builder may be dishonest. Situation- or person-specific beliefs supporting trusting intention quickly become stronger than
dispositional support as the buyer gains experience with the situation.

**Tentative, assumption-based antecedents.** Trusting intention will be fragile at the start of a relationship because of the tentative and assumption-based nature of its antecedents (Figure 2). Initial trust is not based so much on evidence as on lack of contrary evidence (Gambetta, 1988). Illusions of control are almost completely assumption based. If the illusion crumbles, the constructs bolstered by it can decrease rapidly, negatively affecting trusting intention. Riker points out that stereotyping categorization is tentative because "only rarely do [the applied categories] effectively discriminate between the trustworthy and the untrustworthy" (1971: 78). Hence, trusting beliefs (leading to trusting intention) produced by categorization are subject to abrupt corrections. Initial institution-based trust, founded on assumptions about the situation, is subject to rapid deterioration as facts become known. One's disposition to trust assumes things will work out successfully, and will only be salient until situational or personal facts are uncovered (Johnson-George & Swap, 1982).

Experience supplies facts that can quickly displace illusions and assumptions. Fazio and Zanna (1981) point out two reasons for this. First, people consider behavioral experience information to be more reliable than indirectly obtained information. Further, more reliable information reduces uncertainty, making it highly desirable (Smith, Benson, & Curley, 1991). Second, an individual's direct experience in forming an attitude or judgment makes that attitude more readily accessible in his or her memory (cf., Riker, 1971: 78).

**High perceived risk.** Some risk is perceived even when people trust each other (e.g., Mayer et al., 1995). High levels of perceived risk make it more likely that the trustee will pay attention to the other's behavior, seeking validating information. Kramer (1996) found this to be true in relationships between students and faculty members. The more one attends to new information about the other person, the more likely it is that contrary evidence will be found, negatively affecting trusting intention.

**Proposition 10:** In the initial relationship, high trusting intention is likely to be very fragile when it is supported by only one or two antecedents, when it relies almost exclusively on assumptions, and when perceived risk is high.

Several of the above fragility examples demonstrate that the interdependent nature of the model constructs is itself a reason trusting intention may be fragile. When one construct's level is reduced, it is likely to negatively affect a related construct. For example, whereas a decrease in trusting stance directly affects trusting intention, it also affects institution-based trust. Also, while institution-based trust will have a direct effect on trusting intention, it will also affect trusting intention through trusting beliefs. Hence, the model portrays the possibility that trusting intention is like the roof on the proverbial "house of cards," which collapses when one structurally key antecedent slips.

**Conditions Under Which Trusting Intention Is Likely to Be Robust**

Trust these intention is likely to be robust for three reasons: (1) adequate antecedent support, (2) belief-confirming cognitive mechanisms, and (3) social mechanisms.

**Adequate antecedent support.** Given our model, the most obvious reason trusting intention will continue to be high is that, in many cases, several of the antecedents shown in Figure 2 will exist at high levels. Further, it is more probable that the antecedents will exist at a consistently high level rather than in a combination of high and low levels. Researchers of cognitive consistency have found evidence that related beliefs tend to stay consistent with each other because people keep their various cognitions reconciled (Abelson et al., 1968; Luhmann, 1979). Hence, for a given subject, we expect to find relatively consistent levels among trust constructs, especially as the relationship begins. A consistently high set of antecedents will be less likely to cause a trusting intention dissolution than will an inconsistently high set.

For example, in a classroom setting it is likely that, because of past experiences, students feel comfortable with their own and the instructor's roles (institution-based trust). Hence, the students probably have formed structural assurance beliefs within class situations generally. Hopefully, they have heard enough about the instructor's reputation so that, along with early
cues, they can categorize her or him as having trustworthy attributes (e.g., competence), thus forming one or more of the trusting beliefs. Under these circumstances, the students are likely to have high initial trusting intention toward the instructor. This trusting intention probably will prove stable, given it is likely supported by trusting beliefs and institution-based trust. Unless the instructor does something to seriously violate the positive expectations a student develops from the initial class interaction, the student’s trusting intention is likely to endure throughout the course. But trusting intention is even more likely to stay high if the student also has high disposition to trust, which will reinforce his or her institution-based trust and trusting beliefs. A person with a high disposition to trust is more likely to see good points and to overlook flaws in another person or situation that would threaten high trusting intention levels by lowering trusting beliefs or institution-based trust.

Belief-confirming cognitive mechanisms. Attentional cognitive processes play a role in sustaining initial trust. Not all information is attended to, and unless it is, it will not affect trust-related constructs. Indeed, peoples’ beliefs and preconceived notions tend to filter the information they heed. Good (1988) remarks that evidence contrary to one’s beliefs is seldom sought and often ignored, citing psychological experimental evidence (e.g., Mitroff, 1974; Tajfel, 1969; Wason, 1960) for this confirmation bias effect. Similarly, Taylor and Brown cite evidence that people “generally select, interpret, and recall information to be consistent with their prior beliefs or theories (see Fiske & Taylor, 1984; Greenland, 1980; Taylor & Crocker, 1981, for reviews)” (1988: 202). The authors apply this effect to positive illusion-based beliefs: “Consequently, if a person’s prior beliefs are positive, cognitive biases that favor conservatism generally will maintain positive illusions more specifically” (Taylor & Brown, 1988: 202). They cite additional evidence that peoples’ preconceptions guide which information they consider relevant and, therefore, to which information they attend (Howard & Rothbart, 1980; Nisbett & Ross, 1980).

Researchers conducting empirical studies have confirmed that much counterbelief evidence is simply ignored. An example is provided by Starbuck and Milliken’s (1988) study of the 1986 NASA Challenger disaster. The authors found that NASA and Morton-Thiokol managers ignored or explained away evidence that rocket booster O-rings would erode in a low-temperature takeoff. Contrary evidence is especially ignored when people perceive things to be going well. Sitkin, for example, states that “small successes may unintentionally induce low levels of attention and reduced information search” (1992: 232). In contrast, “The greater the incidence of small prior failure, the more attention will be paid to and the more deeply will [be] the processing of information about potential problems” (Sitkin, 1992: 240).

The tendency to ignore counterbelief evidence should be true of trusting beliefs as well as other beliefs. However, in the initial relationship, the desire to feel assured through experiential evidence (Fazio & Zanna, 1981) will, to some extent, mitigate this cognitive bias. This desire will be heightened by low levels of disposition to trust or institution-based trust or by high levels of perceived risk, because the person will require additional assurances that his or her initial beliefs are accurate.

When people do attend to information that disconfirms their views, they often discount it as inaccurate or uninformative (Markus, 1977; Ross & Anderson, 1992; Swann & Read, 1981; Taylor & Brown, 1988), or they reinterpret it positively (Holmes, 1991; Robinson, 1996). Ross and Anderson point out that “beliefs . . . are remarkably resilient in the face of empirical challenges that seem logically devastating” (1982: 144). People tend to accept belief-supportive information uncritically, but slowly acknowledge disconfirming evidence (Tetlock, 1985). Ross and Anderson (1992) posit that people subject disconfirming evidence to considerable scrutiny but do not go back and update or re-evaluate evidence relevant to their beliefs based on disconfirming evidence. People also search their memories to find ways of explaining their existing beliefs (Ross & Anderson, 1982). Similarly, Kahneman and Tversky (1973) found significant evidence that people will interpret ambiguous or incomplete information in a way that agrees with their pre-existing beliefs. If one has a high level of trusting intention toward another, for example, then one can view specific trust violations as isolated exceptions or as a personal quirk (Sitkin & Roth, 1993; Zucker, 1986), with no resulting negative effect on trusting beliefs. High faith in humanity provides another.
reason to ignore or discount such behavior, because it assumes that most people are basically good. Having high faith in humanity will facilitate quick forgiveness of trust violations.

Good (1988) has suggested that another reason for this "cognitive inertia" is the "set effect," which refers to the continued use of mentally stored procedures to handle a situation, even when the situation changes. That is, once people have developed a situational strategy, they tend to continue to use it, even when it does not work. Applied to trust, this means that an individual will continue to trust another, even when the latter breaches the former's trust—at least for a while. Good (1988) uses the set effect to reinterpret some of the results of experimentalists who studied trust with such procedures as the Trucking Game. Note the interactive effect of this cognitive tendency with the model's constructs: high institution-based trust and high disposition to trust jointly encourage a person to believe such a course of action is not very risky.

Luhmann (1979) offers a related reason for why trust may be resilient, stating that people build up mechanisms to handle refutations of their trust defections. This is especially true when the felt security associated with trusting intention is not strong:

Insecure expectations, however paradoxical it may at first appear, are psychologically more stable than secure ones. ... (They are) normalized, stereotyped and thus in various ways immunized against the refutation. Explanations of disappointment are built into them, in such a way that a particular case of disappointment presents no problem but rather confirms the structure of the expectation as a whole" (Luhmann, 1979: 79).

Hence, it appears that people develop mechanisms enabling them to absorb disappointments as part of their expectation of the other person, thus reducing the effect of the disappointment on trusting intention. We speculate that individuals with high levels of disposition to trust will absorb disappointments better than those with low levels.

Proposition 11: In the initial relationship, high trusting intention is likely to be robust when (1) a combination of several of its antecedents encourages the trustor to ignore, rationalize, or absorb the negative actions of the other and (2) continued success or low perceived risk of failure consequences cause little critical attention to be paid to the other's behavior. Subjects with high disposition to trust or institution-based trust levels are likely to pay even less critical attention.

Social mechanisms. Good (1988) notes that being around another person generally will increase an individual's favorable beliefs about that person. This occurs because the interpersonal cues from the other person generally are harder to misconstrue face to face and because the pair can more easily go beyond surface information to more substantive levels of mutual understanding. Hence, high levels of trusting intention likely will be sustained as people interact in cooperative ways. If the individuals hold positive beliefs about each other, they are not likely to decrease interaction (Darley & Fazio, 1980). Therefore, the trusting cycle becomes self-sustaining.

Social interaction also will tend to uphold early trusting intention, because people in social situations tend to confirm their beliefs about themselves (Swann, 1983) and about each other. For example, one who trusts another will tend to express that trust in actions toward the other. Because initially extended trust is usually reciprocated (e.g., Burt & Knez, 1996), the other party will also express trust. This confirms the first party's trusting beliefs, as Dasgupta (1988) notes, which, in turn, supports continued high levels of trusting intention.

Social interaction also upholds early trusting intention because of reputation effects. A person's reputation spreads gradually (Dasgupta, 1988), through social interaction (Burt & Knez, 1996). One party remembers the other party's previous encounters. This interactional history cumulates, along with information about the person's background (Dasgupta, 1988), and is transmitted to others. When many people perceive that an individual has a good reputation, it is harder for a negative event to significantly reduce a high level of trusting beliefs in that individual.

We believe that social interaction also sustains trusting intention through institution-based trust. When parties interact in a cordial way, they establish a feeling and appearance that everything is normal or in proper order (Lewis & Weigert, 1985a). Hence, social interaction sustains situational normality belief.
Through situational normality belief, the parties’ interaction strengthens trusting intention. If one has high levels of situational normality belief, one is also likely to believe that structural assurances will operate properly in the situation. Structural assurance will, in turn, positively influence trusting beliefs and trusting intention, as discussed above.

Proposition 12: In the initial relationship, high trusting intention is likely to be robust when (1) the parties interact face to face, frequently in positive ways, or (2) the trusted party has built a widely known good reputation. Social interaction affects trusting intention by its positive effects on trusting beliefs and institution-based trust.

RESEARCH IMPLICATIONS AND CONCLUSION

Future Research Implications

The interplay between trusting intention and its antecedents helps explain the paradox of high trust in new relationships and situations (e.g., Kramer, 1994), but our model also can help explain “disturbing” results in game theory research (Baxter, 1972: 100)—a well-spring for trust theory historically. In Baxter’s review of the two-person game theory research, he indicates that researchers had found no solid link between trusting personality and trusting behavior in the Prisoner’s Dilemma game. Because our model ties dispositional variables to beliefs and intentions, it suggests that personality variables may be too distal from behaviors to be predictive. Our model’s time boundary suggests that disposition-related trust will only be salient when the parties first meet in an ambiguous situation. Further, the institution-based trust constructs help explain why no solid link was found by game theorists.

Whereas game theorists assume that the game context effectively restricts the effects of situational variables (Baxter, 1972), our model’s use of institution-based trust suggests that what a subject believes about the situation is an empirical question that should be measured as a potential confound (e.g., Erez, 1992), rather than being assumed away. Over 25 years ago, Kee and Knox recommended that Prisoner’s Dilemma research should include “continuous” measures of subjects’ cognitions in addition to their behavioral choices (1970: 365). Very little trust research has been done in this manner. Our model’s constructs provide a reasoned theoretical basis for doing so. We suggest that researchers replicate early behavioral experiments while measuring our constructs. This will enable researchers to reinterpret historical results.

Significant empirical work is needed in order to obtain evidence regarding the propositions we make in this article. First, researchers should develop reliable and valid instruments reflecting the constructs in this article, and they should then use them to test various subsets of the model through questionnaire studies. For example, Propositions 1 and 2 and 4 through 7 could be tested in one study by measuring trusting intention, faith in humanity, trusting stance, structural assurance belief, and benevolence belief.

Second, researchers should test our categorization and illusion propositions in laboratory settings, incorporating disposition to trust and institution-based trust constructs as control variables.

Third, researchers should test the model’s interactive effects. For example, scholars should test whether disposition to trust constructs moderate the effects of categorization on trusting beliefs. Based on Proposition 1, one could test whether or not structural assurance belief moderates the link between faith in humanity and trusting beliefs. Trusting stance and institution-based trust may also have an interactive effect on trusting intention. The combination of a high situational normality belief and a high structural assurance belief would probably produce higher trusting beliefs than either would alone.

Finally, the fragility/robustness of initial trust should be tested, based on our propositions, in which we argued that initial trust may be either fragile or robust under certain conditions. For example, one could test, through longitudinal laboratory experiments, the extent to which social interaction increases the robustness of trusting intention when (1) the relationship is successful/unsuccessful or (2) risk is high or low. Many more conditions could be posited.

Our distinction between trust levels and trust fragility/robustness raises important issues that merit further conceptual and empirical work. For example, a high initial level of trusting intention may be quite fragile (subject to change),
whereas a low initial level of trusting intention may not be very fragile. Thus, an interesting research question is under what conditions are low initial levels of trusting intention less fragile than high initial levels of trusting intention?

Conclusion

Our primary contribution in this article is explaining the high initial trust paradox by synthesizing a model of constructs and processes from diverse trust research streams. Because initial trusting intention is not always high, our model resolves the paradox by pointing out why it could be high initially but may not be high because of the situation or the persons involved. Thus, the model is predictive based upon specified conditions related to the antecedents of trusting intention.

This model, along with our discussion of the conditions producing fragile versus robust initial trust, should generate significant amounts of research. The delineation of two specific types of institution-based trust and disposition to trust constructs provides researchers construct definitions that lend themselves to consistent empirical measurement. By using both higher-level constructs (e.g., trusting beliefs) and lower-level constructs (e.g., honesty belief), we have helped organize the “confusing potpourri” (Shapiro, 1997a: 625) of trust construct definitions researchers must grapple with.

Our model integrates different aspects of trust that have not been linked previously, and it does so within a specific time parameter. Some have called for more integrative models that simultaneously address dispositional and situational constructs (Davis-Blake & Pfeffer, 1999; Sitkin & Pablo, 1992). We have answered this call, creating a model that brings together dispositional, situational, and interpersonal constructs from four divergent research streams.

Poole & Van de Ven (1989) remark that by adding the temporal dimension, one can improve theories. In this article we explain that the processes by which trust forms initially are not the same as those by which it forms later. In particular, a model of continuing trust will emphasize experiential knowledge while de-emphasizing assumptions and dispositions. In this way, the model’s temporal lens highlights the unique aspects of how trust forms at the earliest stage of an organizational relationship.

REFERENCES


Davis, F. D. 1989. Perceived usefulness. perceived ease of


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


**D. Harrison McKnight** is an assistant professor in the College of Business at Florida State University. He received his Ph.D. from the University of Minnesota. His current research interests include organizational trust, ethics, management controls, and the job characteristics of computer personnel.

**Larry L. Cummings** received his Ph.D. in organizational behavior from Indiana University. When Larry passed away in 1997, he was serving as professor of management at the Carlson School of Management, University of Minnesota. In 1995 he was honored by the Academy of Management with the Distinguished Educator Award.

**Norman L. Chervany** is a professor of information and decision sciences at the Carlson School of Management, University of Minnesota. He received his Ph.D. in decision sciences from Indiana University. His current research interests are the adoption of information technology and the effects of information technology on work design.