

# Explicit and Implicit Trust Within Safety Culture

Calvin Burns,<sup>1\*</sup> Kathryn Mearns,<sup>2</sup> and Peter McGeorge<sup>2</sup>

---

Safety culture is an important topic for managers in high-hazard industries because a deficient safety culture has been linked to organizational accidents. Many researchers have argued that trust plays a central role in models of safety culture but trust has rarely been measured in safety culture/climate studies. This article used explicit (direct) and implicit (indirect) measures to assess trust at a UK gas plant. Explicit measures assessed trust by asking workers to consider and state their attitude to attitude objects. Implicit measures assessed trust in a more subtle way by using a priming task that relies on automatic attitude activation. The results show that workers expressed explicit trust for their workmates, supervisors, and senior managers, but only expressed implicit trust for their workmates. The article proposes a model that conceptualizes explicit trust as part of the surface levels of safety culture and implicit trust as part of the deeper levels of safety culture. An unintended finding was the positive relationship between implicit measures of trust and distrust, which suggests that trust and distrust are separate constructs. The article concludes by considering the implications for safety culture and trust and distrust in high-hazard industries.

---

**KEY WORDS:** Distrust; implicit attitudes; safety culture; trust

## 1. INTRODUCTION

A deficient safety culture has been implicated in a number of organizational accidents from the disaster at the Chernobyl nuclear power plant (IAEA, 1991) to, most recently, the loss of the space shuttle Columbia (CAIB, 2003). An organization's safety culture, defined here in terms of Schein's (1990) model as basic underlying assumptions about safety, has been presupposed to affect safety performance through worker safety behaviors and safety management practices (ACSNI, 1993, p. 23). A theoretically important but underinvestigated area involves the role of trust within safety culture. This article investigates the role of trust within safety culture by using explicit (direct)

and implicit (indirect) measures to assess trust at a UK gas plant.

## 2. SAFETY CULTURE AND TRUST

Reason (1997) proposed that an effective safety culture could be socially engineered by developing three subcomponents: a reporting culture, a just culture, and a learning culture. He noted that to create a safety culture, an organization must first have a safety information system but that its effectiveness is dependent on the willing participation of the workforce. To achieve this, he argued that it is necessary to engineer a reporting culture in which employees are prepared to report their errors and near misses but that they are not likely to do so if they think the organization will respond in a punitive manner by blaming and punishing them. Thus, Reason proposed that an organization must have a just culture, in which there is an agreed set of principles for drawing the line

<sup>1</sup> University of Strathclyde, UK.

<sup>2</sup> University of Aberdeen, UK.

\* Address correspondence to Dr. Calvin Burns, Strathclyde Business School, University of Strathclyde, Graham Hills Building, Glasgow, G1 1XU, UK; tel: 44(0)141 548 4251; fax: 44(0)141 552 3581; calvin.burns@strath.ac.uk.

between acceptable and unacceptable actions. This is important because some unsafe acts, like substance abuse, warrant severe sanctions and a no-blame culture (which provides a blanket amnesty on all unsafe acts) would be seen to oppose natural justice. Reason (1997) proposed that in order to create a safety culture, an organization must also possess a learning culture. He argued that this entails the organization having the willingness and the competence to draw the right conclusions from its safety information system, and the will to implement major reforms when their need is indicated. He concluded that a reporting culture, a just culture, and a learning culture interact to create a safety culture.

Reason (1997) argued that the subcomponents of safety culture are based on an underlying element of trust. Although the concept of trust appears in a variety of senses in the social sciences, it is now widely regarded as *a psychological state comprising the intention to accept vulnerability based upon positive expectations of the intentions or behaviour of another* (Rousseau *et al.*, 1998, p. 395). As many authors have shown, certain attitudes and perceptions about an individual can lead to trust in that individual (for a review of the factors of trustworthiness, see Mayer *et al.*, 1995). Thus, this article adopts the conceptualization of trust as a psychological state that can result from the activation of a trust-related attitude for the individual to be trusted.

With respect to engineering a reporting culture, Reason (1997) proposed that for the workforce to report their errors and near misses they must trust managers to treat them fairly. In other words, trust is needed to integrate a reporting culture and a just culture into an effective safety culture. Reason (1997) did not explicitly specify the role of trust in creating a learning culture but he stressed the importance of feedback to sustain incident reporting. Therefore, it is proposed that trust may be important in integrating all three subcomponents into an effective safety culture. Providing workers with rapid and useful feedback after they report incidents should increase their trust in management and reinforce reporting behaviors. This should also enhance understanding of potential hazards and allow the organization to learn from its shortcomings. Other models of safety culture (Hudson, 2003; Westrum, 1995) have also ascribed trust a central role and empirical evidence linking trust and safety performance is starting to emerge (Zacharatos *et al.*, 2005).

### 3. MEASURING SAFETY CULTURE

The most common approach to measuring safety culture is to conduct a safety climate survey of the workforce. These questionnaires assess workers' explicit attitudes and perceptions about aspects of organizational safety (e.g., management's commitment to safety, work pressure, communication), thereby providing indicators of the organization's underlying safety culture. The results of safety climate surveys have been shown to be directly and indirectly linked to safety outcomes such as self-report and officially recorded rates of workforce injuries (Cooper & Phillips, 2004; Donald & Canter, 1994; Lee, 1998; Mearns *et al.*, 1998; Zohar, 2000).

Despite the central role of trust in leading models of safety culture, literature reviews of the factors measured in safety climate surveys (Flin *et al.*, 2000; Guldenmund, 2000) have shown that trust is rarely measured. When questionnaire items about trust are included, it is usually just as one or two items (e.g., Mearns *et al.*, 2000). Moreover, survey instruments may lead to overestimates of trust because they typically give rise to response biases. Researchers interested in prejudice (Dunton & Fazio, 1997; Plant & Devine, 1998) have argued that scores on direct measures not only reflect respondents' attitudes but also the deliberate and conscious manipulation of responses to regulate their impression on others. Thus, survey data about trust may be based on respondents' attempts to convey that they are trusting individuals by stating that they trust their work colleagues more than they actually do.

#### 3.1. Implicit Measures

In an attempt to minimize response biases associated with explicit measures, researchers interested in prejudice have developed techniques that measure attitudes implicitly (i.e., indirectly). Instead of asking for verbal reports, implicit measures rely on the automatic activation of attitudes and the time it takes an individual to react to an attitude-evoking word (see Fazio & Olson, 2003 for a review of implicit measures). On implicit attitude measures like priming tasks, quicker reaction times are indicative of automatic attitude activation whereas longer reaction times represent the lack of automatic attitude activation.

Within the prejudice literature, correlations between explicit and implicit measures tend to be quite

low (e.g., Fazio *et al.*, 1995; Greenwald *et al.*, 1998). Fazio and Olson used the MODE model (Motivation and Opportunity as Determinants of whether the attitude-to-behavior process is primarily spontaneous or deliberative in nature) to explain the discordance between explicit and implicit measures. They stated that *if either motivation or opportunity [to deliberate] is relatively low at the time that the explicit response is being considered, then explicit measures should correlate with implicit ones. However, when both motivation and opportunity are relatively high, they are less likely to correlate* (2003, p. 304). They also used the MODE model to explain the predictive validity of explicit and implicit measures. They stated that *when motivation and/or opportunity are low, behavior is expected to be largely a function of the automatically activated attitude, and hence, the implicit measure should prove predictive. When motivation and opportunity are high, the explicit measure should be more predictive—presumably because the explicit measure will have been influenced by these same motivational forces* (2003, p. 305).

The MODE model offers reasons to support the use of implicit measures in assessing trust within safety culture, apart from the minimization of response bias. This model predicts that implicit attitude measures should influence behaviors in situations when motivation/opportunity to deliberate is low. During high-tempo operations, or situations involving high mental workload, implicit trust (in senior managers, supervisors, and workmates) may lead to spontaneous behaviors like stopping the job for safety reasons, or challenging an unsafe act or rule violation. To further investigate the role of trust within safety culture, this article used explicit and implicit measures to assess trust at a UK gas plant.

#### 4. TRUST AND DISTRUST WITHIN SAFETY CULTURE

Influential models of trust from various academic disciplines have conceptualized trust and distrust as being at opposite ends of a single continuum (Hosmer, 1995; Mayer *et al.*, 1995; Rotter, 1971). For example, Mayer *et al.* (1995) conceptualized high trust, as determined by perceptions of ability, benevolence, and integrity, as indicative of low distrust. However, some theorists have argued that trust and distrust are separate constructs (Deutsch, 1960; Lewicki *et al.*, 1998; Sitkin & Roth, 1993). In Sitkin and Roth's (1993) model, trust rests on an individual's expectations of others' ability to complete tasks reliably, whereas distrust is generated when an individual's values and be-

liefs are not compatible with those of the organization or other individuals in question. Accordingly, Lewicki *et al.* (1998) defined trust in terms of confident positive expectations regarding another's conduct, and distrust in terms of confident negative expectations regarding another's conduct. Psychologists are increasingly calling for reexamination of the assumption that positive-valent and negative-valent attitudes are opposite ends of a single continuum (Cacioppo & Berntson, 1994; Petty *et al.*, 1997). Research findings regarding positive and negative affectivity suggest that these are not opposite ends of the same continuum but exist as distinct bipolar constructs (Burke *et al.*, 1989; Watson & Tellegen, 1985). In order to contribute to this debate, this article used implicit measures to investigate the relationship between trust and distrust.

#### 5. METHODOLOGY

The study consisted of two parts: a questionnaire and a priming task. The questionnaire measured participants' explicit trust in their workmates, supervisors, and senior plant managers, and the priming task measured participants' respective implicit trust. The order of study administration was counterbalanced such that half of the participants completed the questionnaire first and the other half of the participants completed the priming task first. Assignment to the counterbalancing conditions was done randomly and no effects of study order were found. The first author administered the questionnaire and priming task onsite during one-to-one meetings with participants in a private room. Participants were told that the study consisted of two parts and were told if they would start with the questionnaire or the priming task. They were debriefed about the study at the end of both tasks.

##### 5.1. Participants

Fifty-three employees (approximately 50% of the workforce) at a UK gas plant participated in the study. During access negotiations, the authors met with senior management and workforce safety representatives at the gas plant to make a call for voluntary participation by asking the workforce safety representatives to distribute promotional material about the study and to collate a list of volunteers. Participants were randomly selected from volunteers who were onsite during the three-week period in which the study was conducted. Participants were selected from the "sharp end" (i.e., the construction, gas plant services, maintenance and operations job types) because

they are exposed to high levels of risk, thus making trust and safety behaviors critical to them. Of the 53 participants, 21 were employed by the operating company and 32 were employed by one of three different contractor companies. The large percentage of contractor staff is typical in the jobs that were sampled. The absence of female staff is also typical in the jobs that were sampled. Participants' mean work experience in the oil and gas industry was 17.91 years ( $SD = 8.22$ ) and their average work experience at the current plant was 9.12 years ( $SD = 7.20$ ). Participants were given time away from their regular duties to take part in the study.

## 5.2. Questionnaire

The questionnaire items used to measure trust in workmates, supervisors, and the plant leadership were "I trust my workmates," "I trust my supervisor," and "I trust the plant leadership," respectively. All items used a 6-point scale with anchors of "agree" and "disagree" for each scale point. Although it is recognized that trust is context-specific, these items were chosen to measure explicit trust because they are general in nature. This allowed them to be compared with implicit measures of trust, which due to restrictions of the priming task, are also general in nature. The questionnaire contained other items about aspects of organizational safety, but these items are not reported in this article.

### 5.2.1. Procedure

In order to ensure anonymity and confidentiality, no identifying information was collected from participants and they were assured that nobody other than the researchers would see their questionnaires. Participants were told that they could leave a question blank if they did not want to answer it and that they did not need to give a reason for doing so. Participants were told to take as much time as they needed to complete the questionnaire and that they could ask questions at any time.

Participants were given a copy of the questionnaire. They were advised that some of the questions would be about their workmates and supervisor and that in order to avoid confusion, they were to think of their current workmates and supervisor when answering such questions. To remind participants to think of their current workmates, the following statement appeared immediately before items about workmates: "The following questions refer to your current

*Workmates.*" Similarly, the following statement appeared immediately before items about the supervisor: "The following questions refer to your current *Supervisor.*" In the case of contractor employees, the statement read: "The following questions refer to your current *Supervisor* from your *Contractor Company.*"

## 5.3. Priming Task

The procedure for the priming task was adapted from Fazio *et al.*'s (1995) Bona Fide Pipeline to Automatic Activation. This procedure measures the extent to which the presentation of an attitude object (i.e., prime) automatically activates an associated evaluation from memory. In this study, the prime was a job-related word (e.g., supervisor). Its presentation was followed by the display of a trust-related or distrust-related target word. The participants' task was to press a key labeled "trust" or a key labeled "distrust" as quickly as possible to indicate their judgment of the target word. The latency with which this judgment was made constituted the dependent measure.

Suppose, for example, that an individual trusts his workmates. Presentation of "workmates" as the prime may automatically activate an associated trust-related attitude. If the target word that is subsequently presented is also trust related (e.g., loyal), then the individual is able to indicate the connotation of the target word relatively quickly. In this case, responding has been facilitated, which is indicative of implicit trust.

### 5.3.1. Equipment

The priming task was conducted on an IBM compatible laptop computer using Super Lab version 1.03. All visual stimuli were presented on a single line in the center of the screen in bold Times New Roman 60-point font. Responses were made on a 4-button response box with keys labeled "Trust," "Distrust," "Yes," and "No."

### 5.3.2. Development of Stimulus Materials

The target words (words indicative of trust or distrust) used in the task were generated by 50 undergraduate students who were taking part in a group practical exercise for an introductory psychology course. These student participants received course credit for their participation. Collectively, the students generated 182 different words and phrases about trust and distrust. These words and phrases

were later rated for connotations of trust (1 = total lack of trust, 2 = some lack of trust, 3 = not relevant, 4 = trust, 5 = high level of trust) by a different group of 50 undergraduate students during another group practical exercise. The 10 words with the highest trust ratings were selected as targets indicative of trust. These words were: be sure of, caring, confide, count on, dependable, honest, honor, loyal, reliable, truthful. Their mean trust rating was 4.42 ( $SD = 0.26$ ), indicating that they were indicative of trust. The 10 words with the lowest trust ratings were selected as targets indicative of distrust. These words were: backstabber, deceitful, devious, dishonest, double dealing, liar, sly, traitor, twofaced, unreliable. Their mean trust rating was 1.21 ( $SD = 0.19$ ), indicating that they were indicative of a lack of trust. It should be noted that targets were selected so as not to include words with "trust" in them. As expected, a paired samples  $t$ -test revealed that the difference in trust rating between the groups of target words was significant,  $t(49) = 62.05$ ,  $p < 0.001$ . This finding indicates that the trust and distrust targets had very different connotations of trust.

#### 5.4. Procedure

Gas plant participants were made familiar with the laptop computer and 4-button response box, then told that the priming task contained three phases. The three phases were a baseline phase, a priming phase, and a recognition memory test. The purpose of the first phase was to obtain baseline data for the target words. The second phase was the actual priming task. The last phase consisted of a recognition memory test and was included to check whether participants had followed instructions to pay attention to the primes in the priming phase.

The purpose of the first phase was to obtain baseline data for the target words. This phase involved the presentation of a target word on the computer screen. Participants were instructed to press the key labeled "Trust" or the key labeled "Distrust" as quickly as possible to indicate their judgment of the word. Participants performed two blocks of trials. Each block consisted of the randomized presentation of the 20 target words. The average response latency for any given target served as the participant's baseline latency for that target.

The beginning of a trial in the first phase was marked by the presentation of a row of asterisks for 315 milliseconds. Following this, the screen blanked

for 135 milliseconds before the onset of the target word. A given target remained on the screen until the participant responded or for a maximum of 1.75 seconds. A 2.50-second interval separated each trial. Participants' responses and their response latencies (from target onset to key press to the nearest millisecond) were recorded. A block of practice trials preceded this phase. The purpose of the practice block was to familiarize participants with the procedure by using different words than those presented in the first phase.

The second phase involved the actual priming task. Participants were told that they would still have to judge the meaning of target words, but now they would also have to remember different job-related words (primes) that would flash before each target word. Participants were told that it was important for them to pay attention to the primes because they had to complete a recognition memory test in the next phase. Thus, they were led to believe that this phase involved both learning job-related words and judging the connotation of target words.

There were 10 different primes presented during the priming phase. The three critical primes were "workmates," "supervisor," and "plant leadership." The other seven primes served as distracters (e.g., contract company, safety rep). The procedure was similar to the first phase except that the row of asterisks was now replaced by one of the 10 primes. On any given trial, a prime was presented for 315 milliseconds, followed by a 135-millisecond interval before the onset of the target; thus, the stimulus onset asynchrony was 450 milliseconds. Responses and response latencies (measured in milliseconds from target word onset to response) were recorded. A 2.50-second interval separated each trial.

After a short practice block involving different primes and targets, five blocks of trials were presented. Each block consisted of 20 randomized trials, in which each of the primes appeared twice, followed by one of the 20 target words. To ensure that the three critical primes were followed by the same target words over the course of this phase, trust targets were paired randomly with distrust targets to yield 10 pairs of target words. Five pairs of target words were selected randomly and paired with the three critical primes such that each pair of these targets appeared once per block. Thus, over the course of this phase, the same five trust target words and the same five distrust target words followed the three critical primes. The seven distracter primes were paired with pairs of

**Table I.** Primes and Pairs of Target Words Presented in the Priming Phase

Prime	Block				
	1	2	3	4	5
Workmates	Caring	Honest	Loyal	Dependable	Confide
	Traitor	Liar	Dishonest	Backstabber	Unreliable
Supervisor	Honest	Loyal	Dependable	Confide	Caring
	Liar	Dishonest	Backstabber	Unreliable	Traitor
Plant leadership	Loyal	Dependable	Confide	Caring	Honest
	Dishonest	Backstabber	Unreliable	Traitor	Liar
Contract company	Dependable	Truthful	Honor	Count on	Be sure of
	Backstabber	Deceitful	Twofaced	Sly	Devious
Shell	Confide	Honor	Reliable	Honest	Count on
	Unreliable	Twofaced	Double dealing	Liar	Sly
Safety rep	Honor	Reliable	Truthful	Loyal	Dependable
	Twofaced	Double dealing	Deceitful	Dishonest	Backstabber
Maintenance staff	Reliable	Confide	Be sure of	Honor	Loyal
	Double dealing	Unreliable	Devious	Twofaced	Dishonest
Gas plant services	Truthful	Be sure of	Count on	Reliable	Honor
	Deceitful	Devious	Sly	Double dealing	Twofaced
Formal methods sheets	Be sure of	Count on	Caring	Truthful	Reliable
	Devious	Sly	Traitor	Deceitful	Double dealing
Permit to work	Count on	Caring	Honest	Be sure of	Truthful
	Sly	Traitor	Liar	Devious	Deceitful

target words such that each pair of targets only appeared once per block. These pairings are displayed in Table I.

The third phase involved the recognition memory test that participants had been led to expect during the previous phase. The purpose of this phase was to determine whether participants had followed instructions to pay attention to the primes in the previous task. This was important because paying attention to a prime was deemed necessary to automatically activate the associated trust-related attitude. Thus, this phase consisted of the random presentation of the 10 primes used during the previous task along with 10 fillers (new job-related words) that were not presented previously. Participants were instructed to press the key labeled "Yes" if they thought the prime/filler was presented in the previous task, and to press the key labeled "No" if they thought the prime/filler had not been presented earlier. Each prime/filler remained on the screen for a maximum of 5.00 seconds or until the participant pressed a key. A 2.50-second interval separated each trial. Responses were recorded.

## 6. RESULTS

### 6.1. Questionnaire

Participants expressed high levels of explicit trust for their workmates ( $M = 5.03$ ,  $SD = 1.01$ , on a 6-point

scale). They also expressed explicit trust for their supervisors ( $M = 4.82$ ,  $SD = 1.11$ ), and the plant leadership ( $M = 4.27$ ,  $SD = 1.19$ ). Paired samples *t*-tests were conducted in order to test whether the differences between these mean scores were significant. There was a strong trend suggesting that participants trust their workmates more than they trust their supervisor,  $t(52) = 1.98$ ,  $p = 0.05$ . Participants indicated that they trust their workmates more than they trust the plant leadership,  $t(52) = 4.50$ ,  $p < 0.001$ , and that they trust their supervisor more than they trust the plant leadership,  $t(52) = 3.11$ ,  $p < 0.01$ .

### 6.2. Priming Task

#### 6.2.1. Detection Data

During the priming phase, participants were instructed to pay attention to the primes in order to complete a recognition memory test in the next phase. Accordingly, participants completed a recognition task in which they had to indicate whether a given prime or filler had or had not been presented during the priming phase of the task. These data were examined to check whether participants had followed instructions to pay attention to the primes, as doing so was deemed necessary to automatically activate associated trust-related attitudes. Recognition scores were calculated as per the method used by Fazio *et al.*

(1995). For each participant the proportion of false alarms (fillers that the participant incorrectly identified as having been presented previously) was subtracted from the proportion of correctly identified hits (previously presented primes correctly identified as such). Performance at chance levels would be indicated by a score of 0.00 and perfect recognition would be indicated by a score of 1.00. Participants performed at better than chance levels. A one-sample *t*-test indicated that participants' mean score of 0.76 was significantly greater than the chance score of 0;  $t(52) = 27.47, p < 0.001$ .

### 6.2.2. Facilitation Scores

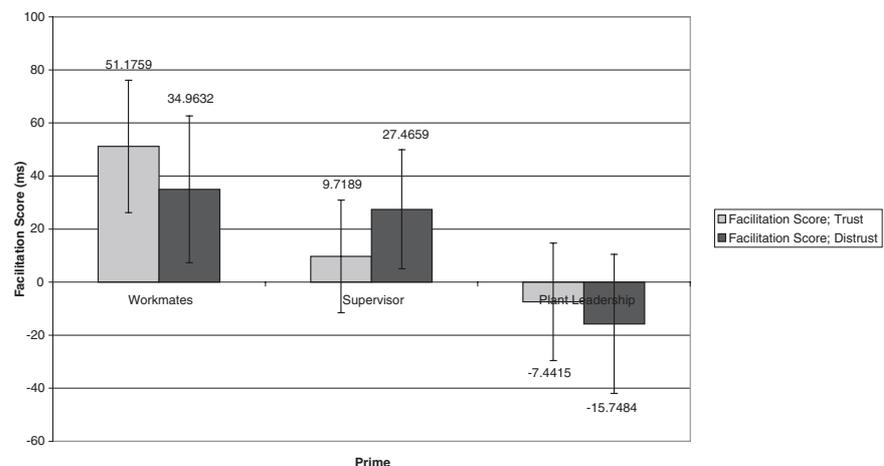
Trust and distrust facilitation scores were computed for each prime to assess the extent to which a prime facilitated responding to the trust and distrust targets. Thus facilitation scores were indicators of implicit trust and implicit distrust.

Facilitation scores were computed as per the method used by Fazio *et al.* (1995). For each participant, baseline latency for each target word was computed from the average of the two presentations of the target word during the first phase of the task. The latency for any given target word when preceded by a given prime was subtracted from the baseline for that word to arrive at a facilitation score. Mean facilitation scores for trust target words and distrust target words were then computed for each prime. The participants' mean facilitation scores are displayed in Fig. 1. The latencies for any trials on which participants made an error or did not respond (mean of 5.53% trials per participant) were not included in these computations.

In Fig. 1, a positive mean trust facilitation score indicated implicit trust. In such cases, the response latencies to the trust targets when preceded by the prime were quicker than the no-prime baselines (i.e., the prime facilitated quicker responding to the trust targets as compared to the no-prime baselines). Conversely, a negative mean trust facilitation score indicated a lack of implicit trust. Here, the response latencies to the trust targets when preceded by the prime were longer than the no-prime baselines (i.e., the prime did not facilitate quicker responding to the trust targets as compared to the no-prime baselines). Mean facilitation scores for distrust were interpreted in the same way but as representing implicit distrust, or a lack thereof, based on the response latencies to the distrust targets.

Paired samples *t*-tests were conducted to test for differences between the primes for the trust and distrust facilitation scores, respectively. Table II displays the results from these analyses.

Table II shows that the trust facilitation score for workmates was greater than that for supervisor and the plant leadership,  $p < 0.05$ , respectively. These findings indicate that participants had more implicit trust for their workmates than they did for their supervisors and the plant leadership. There was a trend suggesting that the distrust facilitation score for workmates was greater than that for the plant leadership,  $p = 0.07$ . Also, the distrust facilitation score for supervisor was greater than that for the plant leadership,  $p = 0.02$ . These findings suggest that participants may have had more implicit distrust for their workmates and supervisors, than for the plant leadership. The differences between the remaining pairs of facilitation scores were not significant.



**Fig. 1.** Mean facilitation scores from the priming task. (Error bar represents the SE of the mean.)

**Table II.** Means, *t*-Statistic, and *p*-Value for Pairs of Facilitation Scores

Pairs of Facilitation Scores	Mean	<i>t</i>	<i>df</i>	<i>p</i>
Trust–Workmates	51.17	2.07	52	0.043
Trust–Supervisor	9.72			
Trust–Workmates	51.18	2.39	52	0.021
Trust–Plant leadership	–7.44			
Trust–Supervisor	9.72	1.03	52	0.307
Trust–Plant leadership	–7.44			
Distrust–Workmates	34.96	0.37	52	0.711
Distrust–Supervisor	27.47			
Distrust–Workmates	34.96	1.85	52	0.070
Distrust–Plant leadership	–15.75			
Distrust–Supervisor	27.47	2.33	52	0.024
Distrust–Plant leadership	–15.75			

To determine the strength of participants' implicit trust and distrust, one-sample *t*-tests were conducted on each of the facilitation scores for each prime to determine whether there was a significant difference from a facilitation score of zero (i.e., no facilitation or baseline latency). The trust facilitation score for workmates was the only facilitation score that was significantly different from zero,  $t(52) = 2.05, p < 0.05$ . These results suggest that participants had strong implicit trust for their workmates but not for their supervisors or the plant leadership. The data also suggest that participants did not have strong implicit distrust for their workmates, supervisors, or the plant leadership.

### 6.2.3. Explicit and Implicit Trust

To investigate the relationship between explicit and implicit trust, correlation analyses were conducted between the questionnaire items about trust for workmates, supervisor, and the plant leadership, and the respective facilitation scores. Table III presents the correlation coefficients and *p*-values from these analyses. No relationship was observed between the explicit and implicit measures of trust for workmates, supervisor, and the plant leadership.

### 6.2.4. Trust and Distrust

This study investigated the relationship between trust and distrust by using implicit measures only. The relationship between explicit measures of trust and distrust was not investigated due to concerns expressed by a manager from the operating company about the length of the questionnaire. Thus, items about distrust were not included in the questionnaire.

**Table III.** Correlation Coefficients and *p*-Values for Trust Facilitation Scores and Questionnaire Items about Trust

Trust Facilitation Score/ Questionnaire Item	<i>n</i>	<i>r</i>	<i>p</i>
Trust–Workmates I trust my workmates	53	–0.02	0.916
Trust–Supervisor I trust my supervisor	53	–0.04	0.755
Trust–Plant leadership I trust the plant leadership	51	0.08	0.571

**Table IV.** Correlation Coefficients and *p*-Values for Trust and Distrust Facilitation Scores

Trust Facilitation Score/ Distrust Facilitation Score	<i>n</i>	<i>r</i>	<i>p</i>
Trust–Workmates Distrust–Workmates	53	0.67	<0.001
Trust–Supervisor Distrust–Supervisor	53	0.58	<0.001
Trust–Plant leadership Distrust–Plant leadership	53	0.44	0.001

Table IV shows the relationships between the trust and distrust facilitation scores for workmates, supervisor, and the plant leadership. Strong positive relationships were observed between the trust and distrust facilitation scores, respectively, for workmates, supervisor, and the plant leadership.

Paired samples *t*-tests were conducted to test for differences between the trust and distrust facilitation scores for each prime. No significant differences were found.

## 7. DISCUSSION

This article investigated the role of trust within safety culture using explicit and implicit measures. The questionnaire found that participants expressed explicit trust for their workmates, supervisors, and the plant leadership. The priming task found that participants only expressed implicit trust for their workmates. In order to investigate the relationship between explicit and implicit trust, correlation analyses were conducted between the questionnaire items about trust for workmates, supervisor, and the plant leadership, and the respective facilitation scores. No relationship was found between these measures (see Table III). These results are consistent with findings reported in the prejudice literature (Fazio *et al.*,

1995; Greenwald *et al.*, 1998) and the predictions made by the MODE model. Participants in this study were given as much time as they needed to complete the questionnaire so opportunity to deliberate about their responses to explicit measures of trust was high, whereas opportunity to deliberate about their responses to implicit measures of trust was low because they responded under speeded conditions. Fazio and Olson (2003) argued that if motivation or opportunity to deliberate is relatively high when the explicit response is being considered, then explicit measures are not likely to correlate with implicit measures, as was the case in this study.

### 7.1. A Model of Explicit and Implicit Trust Within Safety Culture

Reason (1997) proposed that a safety culture could be socially engineered by developing the sub-components of reporting, just, and learning cultures but that trust was needed to integrate these sub-components into an effective safety culture. Schein (1990, 1997) proposed three interconnected levels at which culture manifests itself: observable artifacts, values, and basic underlying assumptions. Although he did not define his basic underlying assumptions in terms of psychological constructs, he argued that it is necessary to understand this category in order to understand the other more surface levels of culture and to predict future behavior.

Implicit attitudes provide one way to extend models of (safety) culture by defining Schein's (1990, 1997) basic underlying assumptions in terms of psychological constructs. Greenwald and Banaji (1995, p. 8) defined implicit attitudes as *introspectively unidentified (or inaccurately identified) traces of past experience that mediate favorable or unfavorable feeling, thought, or action toward social objects*. This definition bears a striking resemblance to *the taken-for-granted, underlying, and usually unconscious assumptions that determine perceptions, thought processes, feelings and behavior* (Schein, 1990, p. 112) that Schein (1997, p. 16) defined as *the essence of culture*. This article defined safety culture as shared basic underlying assumptions about safety, which as per Reason's (1997) model include basic underlying assumptions about trust, reporting, punishment, and organizational learning. Thus, a model of trust within safety culture is proposed to be that implicit trust comprises some of the basic underlying assumptions that are the essence of safety culture, while explicit trust is part of the surface levels of safety culture. The role of explicit trust in the pro-

posed model is consistent with Denison's (1996) argument that explicit measures are restricted to the assessment of organizational climate, which Schein considers *only a surface manifestation of culture* (Schein, 1990, p. 109).

To provide evidence to support the proposed model, one must examine safety outcomes such as the number of accidents, trips to the sickbay, and days absent, at the level of the organization (Hackman, 2003). An organization with a more positive safety culture should exhibit fewer negative safety outcomes and more positive responses on both explicit and implicit measures of trust than another organization operating within the same sector. Thus, in a good safety culture, explicit and implicit trust should be expressed for all members of the organizational hierarchy. With respect to the proposed model, the results from this study are not necessarily indicative of a good safety culture. Certainly, however, the proposed model provides a framework for measuring the basic underlying assumptions about trust that are part of the essence of safety culture and this may prove useful in managing attitude and cultural change within safety critical organizations.

### 7.2. Trust and Distrust Within Safety Culture

If trust and distrust are opposite ends of the same continuum, then strong inverse relationships should emerge between measures of trust and distrust. In order to test this hypothesis, correlation analyses were conducted between the trust and distrust facilitation scores for workmates, supervisors, and the plant leadership. Strong positive relationships were found between the trust and distrust facilitation scores (see Table IV). These data support models of trust (Deutsch, 1960; Lewicki *et al.*, 1998; Sitkin & Roth, 1993) that have conceptualized trust and distrust as separate constructs. Moreover, these data are consistent with other findings in the trust literature (Clark & Payne, 1997; Robinson *et al.*, 1991) that suggest that trust and distrust are separate constructs.

The existence of trust and distrust as separate constructs allows for the possibility of ambivalence or the coexistence of trust and distrust toward the same attitude object (Lewicki *et al.*, 1998). Quasi-experimental research on ambivalence has shown that attitudes about trust and distrust can coexist, and that this is a condition that people intuitively understand (Priester & Petty, 1996). The results of the priming task support the notion of ambivalence and have important implications for trust and distrust within safety culture.

In terms of Lewicki *et al.*'s (1998) model of trust and distrust, the data from the priming task suggest that employees' relationships with their workmates are characterized by high trust and high distrust. The trust facilitation score for workmates was significantly greater than zero though not significantly greater than the distrust facilitation score. Although the distrust facilitation score for workmates was not significantly greater than zero, in the absence of markers, this pattern of results is more suggestive of high trust and high distrust than of high trust and low distrust.

Lewicki *et al.* (1998) proposed that under conditions of high trust and high distrust, one party has reason to be highly confident in another in certain respects, but also has reason to be strongly wary and suspicious in other respects. Many positive experiences, which were trust reinforcing, and many negative experiences, which were distrust reinforcing, are likely to have occurred during the relationship. Lewicki *et al.* (1998) concluded that in modern organizations, relationships based on trust and distrust emerge as the most prevalent form as relationships mature and interdependencies are expanded. In order to create the more optimal conditions of high trust and lower (but not low) distrust for workmates, steps need to be taken to limit those experiences engendering distrust. An example of this with respect to safety could include approaching one's workmates about an unsafe act or rule violation in private, so as not to demean the workmates in the company of others.

The current data also suggest that employees' relationships with their supervisors and the plant leadership are characterized by low trust and low distrust. The trust facilitation scores for supervisor and the plant leadership were not greater than zero. The distrust facilitation score for supervisor was significantly greater than the distrust facilitation score for the plant leadership but neither of these scores were significantly greater than zero.

Lewicki *et al.* (1998) suggested that under conditions of low trust and low distrust, an individual has neither reason to be confident nor reason to be wary and watchful. The authors proposed that such relationships are likely characterized by limited experiences and interaction. With respect to models of safety culture (Hudson, 2003; Reason, 1997; Westrum, 1995), relationships based on low trust are not conducive to propelling the system toward being increasingly informed (these authors argued that an informed culture is a more mature safety culture). Supervisors and the plant leadership need to demonstrate their commitment to safety to the workforce

directly, by being seen to take rapid action when incident reports are made and engaging the workforce in meaningful discussions about safety on the shop floor. Doing so should, over time, give rise to trust for supervisors and the plant leadership. Moreover, by treating the workforce fairly when incidents occur, supervisors and the plant leadership may limit the workforce's negative expectations about their conduct that may lead to distrust. Through consistent and repeated application of the above recommendations, this situation has the potential to change from low trust and low distrust to the more optimal high trust and low distrust.

This section has provided evidence that suggests that trust and distrust are separate constructs, and that trust and distrust can exist simultaneously for work colleagues. In light of these findings, models of safety culture (Hudson, 2003; Reason, 1997; Westrum, 1995) that have conceptualized trust and distrust as opposite ends of the same continuum should be reconceptualized in order to make clear that trust and distrust are separate constructs. Reconceptualizing trust and distrust as separate constructs should allow for the development of more sophisticated models of safety culture by clarifying the role of distrust.

### 7.3. Limitations

It is appropriate at this point to acknowledge some limitations about using facilitation scores as measures of implicit trust/distrust. As facilitation scores are based on reaction times, there are no absolute end points as on a Likert-type scale. A facilitation score that is significantly greater than zero should be regarded as being indicative of strong implicit trust/distrust but from this study it remains unclear just how strong this might be. In order to provide markers for comparison, future studies should collect implicit trust/distrust data for presumably neutral primes like "postman" and "butcher," presumably trusted primes like "priest" and "doctor," and presumably distrusted primes like "politician." These markers should be collected independently of any field studies. Including these markers as distracter primes in future field research could make critical primes like "workmates," and "supervisor" more salient and affect the results.

Another issue concerns variability in the mean trust and distrust facilitation scores. A likely explanation for the large standard errors of the means (see Fig. 1) is that participants may have lost their concentration on a few trials due to the length of each block in the priming phase. This seems plausible given that

there were 20 trials per block and that participants could only rest between blocks. Thus, if a participant experienced a momentary lapse of attention (even as small as 100 milliseconds, one-tenth of a second) on just one of the five trials for a given prime, his mean facilitation score would be artificially decreased. This would result in a larger standard error of the mean when the mean for that facilitation score was calculated for all participants. Because of the possibility of being distracted during the priming task, it was difficult to determine whether a potential outlier was in fact an outlier or a facilitation score reflecting less implicit trust/distrust for a prime. Moreover, removing say two outliers from the calculation of the mean facilitation score for a given prime for a majority of participants would raise the question of validity. The resulting mean facilitation score would then only be based on three responses to target words and the set of target words to which the responses would be based would presumably be different for each participant. Therefore, the data were analyzed without removing potential outliers. In an attempt to remove the influence of outliers by other means, the data were also calculated using median facilitation scores instead of mean facilitation scores but there was no change in the results.

## 8. CONCLUSION

Safety culture is an important topic for managers in high-hazard industries because a deficient safety culture has been linked to organizational accidents. This article argued that trust plays a central role in models of safety culture and proposed a model based on explicit and implicit trust. A call was also made to reconceptualize models of safety culture to make clear that trust and distrust are separate constructs in order to develop the role of distrust.

## ACKNOWLEDGMENTS

The authors would like to thank two anonymous reviewers for their valuable comments on an earlier draft of this article, and the organization that allowed access to its personnel in order to conduct the study. This organization chose to remain anonymous.

## REFERENCES

ACSNI. (1993). *Organising for Safety* (advisory committee for the safety of nuclear installations: Human factors study group third report). Sheffield: HSE Books.

- Burke, M. J., Brief, A. P., George, J. M., Robertson, L., & Webster, J. (1989). Measuring affect at work: Confirmatory analyses of competing mood structures with conceptual linkage to cortical regulatory systems. *Journal of Personality and Social Psychology*, *75*, 1091–1102.
- Cacioppo, J. T., & Berntson, G. G. (1994). Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates. *Psychological Bulletin*, *115*, 401–423.
- CAIB. (2003). *Columbia Accident Investigation Board Report*. Washington, DC: NASA.
- Clark, M. C., & Payne, R. L. (1997). The nature and structure of workers' trust in management. *Journal of Organisational Behavior*, *18*, 205–224.
- Cooper, M. D., & Phillips, R. A. (2004). Exploratory analysis of the safety climate and safety behaviour relationship. *Journal of Safety Research*, *35*, 497–512.
- Denison, D. R. (1996). What is the difference between organizational culture and organizational climate. A native's point of view on a decade of paradigm wars. *Academy of Management Review*, *21*, 619–654.
- Deutsch, M. (1960). The effect of motivational orientation upon trust and suspicion. *Human Relations*, *13*, 123–139.
- Donald, I., & Canter, D. (1994). Employee attitudes and safety in the chemical industry. *Journal of Loss Prevention in the Process Industries*, *7*, 203–208.
- Dunton, B. C., & Fazio, R. H. (1997). An individual difference measure of motivation to control prejudices reactions. *Personality and Social Psychology Bulletin*, *23*, 316–326.
- Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology*, *69*, 1013–1027.
- Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology*, *54*, 297–327.
- Flin, R., Mearns, K., O'Connor, P., & Bryden, R. (2000). Measuring safety climate: Identifying the common features. *Safety Science*, *34*, 177–192.
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem and stereotypes. *Psychological Review*, *102*, 4–27.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*, 1464–1480.
- Guldenmund, F. W. (2000). The nature of safety culture: A review of theory and research. *Safety Science*, *34*, 215–257.
- Hackman, J. R. (2003). Learning more by crossing levels: Evidence from airplanes, hospitals and orchestras. *Journal of Organizational Behavior*, *24*, 905–922.
- Hosmer, L. T. (1995). Trust: The connecting link between organizational theory and philosophical ethics. *Academy of Management Review*, *20*, 379–403.
- Hudson, P. (2003). Applying the lessons of high risk industries to health care. *Quality and Safety in Health Care*, *12*(Suppl 1), i7–i12.
- IAEA. (1991). *Safety Culture* (International Safety Advisory Group, Safety-Series 75-INSAG-4). Vienna: International Atomic Energy Agency.
- Lee, T. (1998). Assessment of safety culture at a nuclear reprocessing plant. *Work & Stress*, *12*, 217–237.
- Lewicki, R., McAllister, D., & Bies, R. (1998). Trust and distrust: New relationships and realities. *Academy of Management Review*, *23*, 438–458.
- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, *20*, 709–734.

- Mearns, K., Flin, R., Gordon, R., & Fleming, M. (1998). Measuring safety climate on offshore installations. *Work & Stress, 12*, 238–254.
- Mearns, K., Whitaker, S., Flin, R., Gordon, R., & O'Connor, P. (2000). *Benchmarking Human and Organisational Factors in Offshore Safety* (Report No. 2000 036). London: HSE OTO.
- Petty, R. E., Wegener, D. T., & Fabrigar, L. R. (1997). Attitudes and attitude change. *Annual Review of Psychology, 48*, 609–647.
- Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology, 75*, 811–832.
- Priester, J. R., & Petty, R. E. (1996). The gradual threshold model of ambivalence: Relating the positive and negative bases of attitudes to subjective ambivalence. *Journal of Personality and Social Psychology, 71*, 431–449.
- Reason, J. (1997). *Managing the Risks of Organizational Accidents*. Aldershot: Ashgate.
- Robinson, J. P., Shaver, P. R., & Wrightsman, L. S. (1991). *Measures of Personality and Social Psychology Attitudes*. San Diego: Academic Press.
- Rotter, J. B. (1971). Generalized expectancies for interpersonal trust. *American Psychologist, 35*, 1–7.
- Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *Academy of Management Review, 23*, 393–404.
- Schein, E. (1990). Organizational culture. *American Psychologist, 45*, 109–119.
- Schein, E. (1997). *Organizational Culture and Leadership*. San Francisco: Jossey-Bass.
- Sitkin, S., & Roth, N. (1993). Explaining the limited effectiveness of legalistic “remedies” for trust / distrust. *Organization Science, 4*, 367–392.
- Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin, 98*, 219–235.
- Westrum, R. (1995). Organisational dynamics and safety. In N. McDonald, N. Johnston, & R. Fuller (Eds.), *Applications of Psychology to the Aviation System*. Aldershot: Avebury.
- Zacharatos, A., Barling, J., & Iverson, R. (2005). High-performance work systems and occupational safety. *Journal of Applied Psychology, 90*, 77–93.
- Zohar, D. (2000). A group-level model of safety climate: Testing the effect of group climate on microaccidents in manufacturing jobs. *Journal of Applied Psychology, 85*, 587–596.