Grant Title:	Measuring Mistrust in Healthcare
Principal Investigator:	Thomas A. LaVeist, Ph.D. Professor of Health Policy and Management
Organization:	Johns Hopkins Bloomberg School of Public Health
Inclusive Dates of Project:	7/1/2002 - 12/31/2003
Federal Project Officer:	JUDY SANGL
Grant Award Number:	R03 HS13274

Structured Abstract

Purpose:

To determine the reliability and validity of a new measure of patient's trust of medical care. The measure is a 17-item scale with three subscales.

Scope:

This study is a community survey of Baltimore, Maryland, via interview.

Methods:

We conducted a telephone survey of a random sample of residents of Baltimore City, Maryland, and interviewed a household member age 18 or older. We surveyed 401 persons and followed up with 327 (81.5%) people 3 weeks after the baseline.

Results:

Factor analysis revealed three subscales that successfully measure the competency, control, and agency dimensions of trust (Chronbach's alpha ranged from .61-.68). However, the internal consistency of the full scale is stronger than that of the subscales (alpha=.77). Test-retest reliability measures for the full scale and the subscales were moderately strong, ranging from .57 to .72. The measure was significantly correlated with the Generalized Trust Scale (corr=.191; $p<.0001$) and the Trust in the Physician Scale (corr=.27; $p<.0001$). The Medical Mistrust Index is a reliable measure of individual's trust of medical care.

Key Words:

Trust, Mistrust, Distrust, measurement, multidimentional scaling, relability, validity, survey research, research methods

Purpose

This application is in response to Program Announcement PA-02-072, "Methodology and Measurement in the Behavioral and Social Sciences." We are requesting support for a study to determine the psychometric properties of a new measure of trust/mistrust of the healthcare system and its performance as a determinant of utilization of preventive and curative care. The PI has a particular interest in its application in dental health services utilization. Dental preventive health services are discretionary and, as such, may be more greatly influenced by psychosocial factors. The Medical Mistrust Index Version 2.2 (MMIv2.2) was initially developed to test the hypothesis that generalized mistrust of the medical care system is a determinant of under-utilization of medical care services among racial minorities and, as such, will help explain race disparities in utilization of healthcare services. However, we believe that the concept of trust/mistrust has much broader implications for healthcare research. Development of the measure has been ongoing since 1996. In a pilot study using a small "convenience sample," the MMIv2.2 has displayed promising psychometric properties (as will be demonstrated below). We now seek support to conduct a study to validate the measure and establish its psychometric properties in a systematic sample. There are two specific aims:

Specific Aim 1: To establish the reliability of the Medical Mistrust Index Version 2.2

Specific Aim 2: To establish the validity of the Medical Mistrust Index Version 2.2

Scope

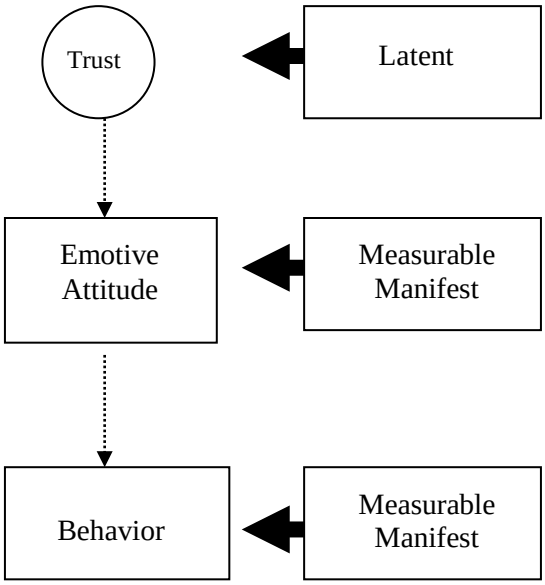
The Trust in Physicians Scale (TIPS) is the most widely used measure of trust within medical care settings. It assesses patients' levels of trust of their individual physicians (Anderson and Dedrick 1990; Thom, et al. 1999B). However, over the past decade, there have been important changes in the dominant modes of medical care delivery (Mechanic 1996). Patient's interactions with the medical care system have become less focused on the individual physician. Increasingly, the patient-provider relationship is with an organization rather than with an individual physician (Mechanic and Schlesinger 1996). Also, low-income and minority patients are more likely to rely on clinics or emergency rooms as their usual source of care. Additionally, some preventive health services, such as mammography or even flu shots, are typically conducted by a technician with whom the patient does not have an ongoing relationship.

Trust is the belief that one will not exploit the vulnerability of another (Sable 1993). But on what basis does one come to this belief? Worchel et al. (1979) has identified three theoretical perspectives on trust. One view is that trust is an emotive response to an individual or other entity. For example, after one brief encounter with a physician, a patient might come to believe that this doctor will

act in the patient's best interests, thus leading the patient to accept the doctor's advice. A second conceptualization is that trust is an experiential phenomenon. Trust is developed over time and multiple experiences, after which one might come to believe that an individual is deserving of trust (Gambetta 1988, Zucker 1986). A third perspective is that trust results from a calculated decision to believe that, within an exchange relationship, an individual/entity will not exploit another's vulnerabilities (McAllister 1995, Rotter 1980). In the absence of prior experiences, this form of trust might come from the reputation of the object of trust. For example, the calculated decision to believe that a highly respected "brand" such as the Mayo Clinic (a Minnesota-based healthcare facility) would not employ unskilled doctors at the Mayo Clinic facility in Florida. The Florida staff, while not the same individuals who earned the high level of public trust and goodwill that the Minnesota Mayo Clinic enjoys, would benefit from that trusted reputation by virtue of having access to the Mayo Clinic name.

These three perspectives on trust suggest that trust has an emotional or attitudinal component in addition to a behavioral component. The measurement model outlined in Figure 1 displays this. Trust is viewed as an unobserved latent construct that can be operationalized only by manifest indicators. The most common manifest indicators (and the two that are relevant for healthcare) are attitude expression and behavior. The Medical Mistrust Index measures the attitudinal component. The dashed lines between the latent construct and the manifest indicators recognizes that there is measurement error between the constructs. Thus, if one has trust (or mistrust), one would be expected to express it attitudinally, but one would expect that this attitude would be manifested by behavioral indicators, such as keeping appointments or adherence to medical advice.

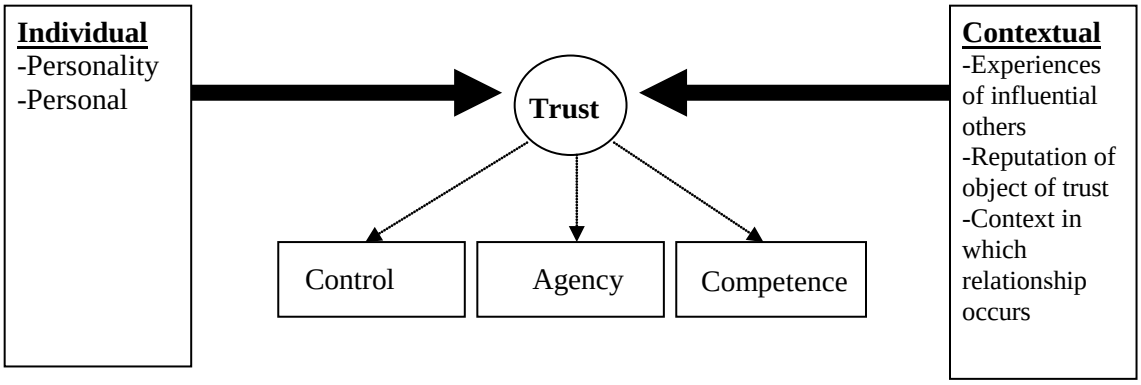
Figure 1. Measurement Model of Trust



In addition to Worchel's et al. (1979) three theoretical perspectives on trust, Mechanic and Schlesinger (1996) identified three dimensions of trust: competence, control, and agency. Competence refers to the expectation that healthcare providers are adequately trained and will perform their responsibilities in a technically proficient way. Control refers to the belief that those who are entrusted with their care will assume responsibility for them and not inappropriately defer to the judgment of others. Agency refers to the proposition that other competing interests will supercede the interests of the patient. The agency dimension is particularly problematic, as physicians increasingly have competing allegiances to their patients and to insurers (Shortell, et al. 1998; Sleeper et al. 1998). Patients primarily are concerned with staying healthy, getting better, living with illness, or coping with the end of life (Institute of Medicine 2001). However, insurers are interested in maximizing cost efficiency (Flood 1998). These divergent interests sometimes clash, and an erosion of patient's trust in the agency role of physicians seems a likely result.

Our conceptual model of trust combines the three theoretical perspectives identified by Worchel et al. (1979) with the three dimensions outlined by Mechanic and Schlesinger (1996). Figure 2 displays the conceptual model of trust. The Figure shows that trust results from a combination of individual-level characteristics of the patient and contextual factors that define the parameters of the doctor-patient relationship. Among the individual-level characteristics are the patient's trait level of trust, which we conceptualize as a characteristic of the individual's personality. Some individuals have a greater predilection toward trust than others. Another individual level factor is the prior experiences of the individual. These experiences may be direct – as in prior experiences with the specific doctor – or they may be indirect experiences – such as prior experiences with doctors, dentists, or hospitals in general. Trust also is influenced by the individual's reaction to characteristics of the object of trust. For example, patients may be more prone to trust a physician of their same racial/ ethnic group, gender, age, religious orientation, or some other characteristic.

Figure 2. Conceptual Model of Trust



Among the contextual factors influencing trust are the prior direct or indirect experiences of influential others with the object of trust. If friends, relatives, or others have had positive experiences, then trust might be transferred: trust bestowed by proxy. This process of transferring trust may also occur by virtue of the reputation of the object of trust or the reputation of the influential other who is vouching for the object of trust.

These individual and contextual influences may be manifested in varying levels of the three dimensions of trust (i.e., control, agency, and competence). A patient may have a high level of trust that the medical staff at a given facility is skilled (high competence trust). But, it is possible to simultaneously believe that competing priorities would influence the staff to make decisions that are not in the patient's best interests (low agency trust). One could also envision scenarios whereby patients could hold divergent attitudes on the different dimensions of trust. This necessitates the creation of a multi-dimensional measure of trust.

Methods

We conducted a series of focus groups among White and African-American patients from a Baltimore-based community hospital. One of the objectives of the focus groups was to generate hypotheses regarding non-financial barriers to utilization of health. The focus group participants were probed regarding the various theoretical dimensions of trust. The transcripts of these focus groups were distilled into a set of items that formed the Medical Mistrust Index (MMI). The items were formatted to follow the format used in the Cultural Mistrust Index (Terrell, et al. 1993). The results of an examination of the psychometric properties of an earlier version of the measure were presented at the 2001 annual meetings of Academy Health (LaVeist et al. 2001). Although the results regarding reliability were promising, we decided to continue development of the measure, because results regarding construct validity were not as expected. Some questions were reworded, and new items were created to conform more closely to the theoretical model described above (the previous version had only one domain). In the present analysis, we examine the reliability and validity of a revised three-domain version of the measure, MMIv2.1.

Reliability

Reliability refers to the stability and/or equivalence of a measure (Aday 1989). We will examine two forms of scale reliability for the MMI, internal consistency reliability, and test-retest reliability. Analysis of the internal consistency reliability of the MMI will include the Cronbach's alpha coefficient, item-to-total correlation, and factor analysis. These tests were computed for the entire 17-item scale and for the three individual subscales. We also will perform factor analysis (principle components analysis) to determine whether the overall scale conforms to the three dimensions of trust that we are attempting to measure (agency, control, and competence).

Test-retest reliability reflects the degree of correspondence between the same respondent's answers to the same set of questions at two or more time points. Internal consistency reliability refers to the consistency or equivalence of the individual items intended to measure trust. If the MMI is internally consistent, we would expect that the individual items that comprise the scale would be correlated.

Validity

Validity refers to the degree to which there are systematic differences between the information obtained in response to the questions relative to the meaning of the concept they were intended to, or related questions about the same concept. Because there is no exact criterion measure, we used two related measures, the Trust in Physicians Scale (Anderson and Dedrick 1990, Thom et al. 1999) and the Generalized Trust Scale, a subscale of the Trust Inventory (Couch et al. 1996). We anticipate that there will be a correlation between MMI and the other trust measures.

Study Sample

There are several important considerations involved in determining the appropriate sample size for this study. One consideration is how to account for sample attrition between wave 1 and wave 2 of the survey. A second concern is allowing for tests of the structural stability of the measure within race- and gender-specific subgroups. A third consideration is accounting for design effects resulting from clustering. We propose to use the Mitofsky-Waksberg (Waksberg 1978) method for sampling households. This method offers advantages in terms of efficiency and costs. However, respondents are sampled within primary sampling units that are associated with geographic location, so the respondents are geographically clustered. Thus, sample size computations must account for this clustering.

Sample size was estimated based on the assumption that medical mistrust will have a standard deviation of 14.2 (based on estimates from Thom et al. 1999b) and a two-tailed 95% confidence interval, alpha = .05. The formula is:

$$[1] \quad n = \frac{z^2 \cdot \Phi}{d^2}$$

$$[2] \quad 123.9 = \frac{(1.96)^2 \cdot 201.64}{(2.5)^2}$$

Thus, after rounding to the nearest whole number, we estimate that we will need 124 respondents at wave 2 of the survey. Assuming 25% sample attrition, we will need 165 respondents for the first wave of the survey.

Because the sampling method leads to clustering of respondents within PSUs, we will include an inflation factor to the sample estimate to account for this design effect (Campbell et al. 2001). Respondents living with geographic clusters may have certain commonalities that might predispose them to certain responses to the MMIv2.1 questions. For example there may have been an incident at a local healthcare facility that influenced respondents living near that facility. If this is the case, the responses of the individual survey respondents would not be independent. Rather, they would be influenced by exposure to a common second-order factor. Consequently, variances among these respondents would be constrained resulting in an underestimated standard error. A consequence of the small standard error is a greater likelihood of type 1 error (research findings incorrectly indicating a significant effect). Simpson et al. (1995) found that fewer than 60% of published studies that they examined (all of which used cluster randomized designs) took clustering into account. This finding was consistent with Donner et al. (1990).

To adjust for this potential problem and achieve the equivalent power of a standard simple random sample, sample size estimates must be inflated by a design effect (Campbell et al. 2001). Following Bland (2000), the design effect can be computed as: 1+(m-1)ICC, in which m is the average number of individuals per cluster and ICC is the intracluster correlation coefficient. Using Bland's conservative estimate of .05 for the ICC and assuming an average of 25 respondents per cluster, we compute the design effect as follows: 1+(10-1).05=1.45. Thus, our standard sample size estimate of 165 must be inflated by 45%, which is 165(1.45)=239.25.

In order to test for differences within sex and race groups, we will need sufficient power to detect differences in proportions between two group. In Table 1, we present the estimated sample sizes needed to detect proportion differences by gender. Based on results from the first version of the MMI, a typical proportion of men reporting mistrust is 25%. Assuming 80% power and an alpha of .05, the table shows samples sized needed to detect gender differences. The table shows that the sample size of 239 would be able to detect each proposed difference except for 35% female and 25% male. To detect this difference, we will need a sample size of 367.

As a result of these findings, I decided to determine the feasibility of obtaining a sample size of 367 over a 7-month field period. By interviewing 15 respondents per week for 25 weeks (slightly over 7 months), we could obtain 375 interviews.

We conducted a telephone survey of a random sample of residents of Baltimore City, Maryland. We sampled households and selected the household member age 18 or older who had the most recent birthday. Baltimore City has 167 telephone exchanges (first thre numbers of a telephone number) within two area codes 410 and 443. The 45 exchanges that were associated exclusively with cellular phones were excluded.

Another 23 exchanges were excluded because they were owned exclusively by large businesses or institutions, such as universities, large corporations, or city and state governments.

Table 1. Summary of Sample Size Power Analysis, 80% Power, Alpha .05	
Male Sample Proportions	Female Proportion 25%
35%	367
40%	178
45%	107
50%	89
55%	65
60%	49

The remaining 99 exchanges were entered into SPSS with all possible combinations of the last four digits (0001-9999). This generated a sampling frame of 989,901 telephone phone numbers. We selected a one percent random sample (9,899). Trained interviewers called each number, documenting those that were disconnected or not in service, those who did not speak English, those who refused, and those who accepted to do the baseline interview. For the telephone numbers answered by an answering machine, a message was left and each number was called back a minimum of two times. The interviewers made contact (actually talked with an eligible respondent) with 783 people; 401 completed the baseline interview (51.2%), and 382 refused.

The interviewers obtained oral consent. The average baseline interview lasted approximately 15 minutes. Participants were told that they would be called back in approximately 3 weeks, and a convenient time to call and some appointments were made to facilitate callbacks. Of the 401 completed baseline interviews, 327 (81.5%) completed the 2nd wave interview. The 2nd wave interview also was done over the telephone and lasted approximately 12 minutes. Respondents were compensated $20 for their participation.

Results

The sample has a mean age of 47.3 years, with respondents evenly distributed across each age category. Nearly 15% of the sample was under age 25, and 19.6% were over age 65. The modal age category was 45-54 years. Men represented 28% of the sample. The sample reflects the race distribution of Baltimore, MD: 69% were African American, 25% were White, and 5.7% were Hispanic or Asian American. Nearly 24% reported incomes below $10,000, and 25% reported incomes above $50,000. Just below 75% of respondents had at least a high school education, and 22.1% were college graduates. Nearly 51% had private insurance, and nearly 24% were uninsured.

Table 1. Demographic profile of the sample.		
Variable		Percent
Age	Younger than 25	14.9
	25-34	12.1
	35-44	17.4
	45-54	21.7
	55-64	14.4
	65 or older	19.6
Sex	Female	28.7
	Male	71.3
Race	White	25.2
	Black	69.1
	Other	5.7
Income	Less than $5,000	8.7
	$5000-$9999	15.2
	$10,000-$14,999	11.5
	$15,000-$24,999	12.9
	$25,000-34,999	13.8
	$35,000-49,999	12.9
	$50,000-$59,000	8.1
	$60,000 or more	16.9
Education	Less than high school graduate	25.6
	High school graduate	32.7
	Some college	19.6
	College graduate	22.1
Health Insurance	Medicaid	21.9
	Medicare	32.9
	Private	50.9
	Uninsured	23.9

In the first set of analyses, we assessed internal consistency. The response categories for the items were (1) strongly agree, (2) agree, (3) disagree, and (4) strongly disagree. Items 5 and 7-11 were reverse coded so that, for each item, a higher score indicates greater mistrust. We conducted principal components analysis with Varimax rotation. This analysis resulted in a three-factor solution, that is displayed in Table 2. Items that did not load on a factor at .5 or greater were dropped from the scale. This resulted in an 11-item final scale. Attempts to further reduce the number of items in the scale resulted in significant degradation of the scale's psychometric properties.

Table 2. Assessment of internal consistency.

	Principal Components Analysis, Varimax Rotation			Item characteristics		
Question	Control	Agency	Competency	Mean	Std. Dev	Item-scale correlation
1) You'd better be cautious when dealing with healthcare organizations.	.57	.07	.14	2.93	.66	.54 p<.0001
2) Patients have sometimes been deceived or mislead by healthcare organizations.	.58	.29	.17	2.70	.66	.65 p<.0001
3) Healthcare organizations often want to know more about your business than they need to know.	.70	-.08	.08	2.48	.70	.51 p<.0001
4) When healthcare organizations make mistakes, they usually cover it up.	.54	.39	.07	2.63	.66	.59 p<.0001
5) The patient's medical needs come before other considerations at healthcare organizations. [1]	-.08	.58	.37	2.28	.64	.50 p<.0001
6) Healthcare organizations are more concerned about making money than taking care of people.	.40	.69	.08	2.56	.68	.66 p<.0001
7) Healthcare organizations put the patient's health first. [1]	.19	.74	.35	2.31	.55	.54 p<.0001
8) Patients should always follow the advice given to them at healthcare organizations. [1]	.04	.15	.60	2.25	.60	.46 p<.0001
9) I trust that healthcare organizations check their staff's credentials to make sure they are hiring the best people. [1]	.15	-.16	.71	2.21	.58	.45 p<.0001
10) They know what they are doing at healthcare organizations. [1]	.35	.16	.57	2.28	.53	.59 p<.0001
11) I trust that healthcare organizations keep up with the latest medical information.[1]	.01	.27	.62	2.05	.49	.50 p<.0001
Cronbach's alpha	.68	.68	.61			

[1] Item was reverse coded

Four of the five items theorized to load on the "control" dimension loaded above .5. Three of the five items theorized to load on the "agency" dimension loaded above .5, and four of seven items loaded on the "competency" dimension as theorized. We computed the subscales and the full scale by summing responses across the items. Cronbach's alpha for the subscales ranged among .61, .68 and .68 for the control, agency, and competency subscales, respectively. The alpha for the full scale was .77. All item-scale correlations were between .45 and .66.

In Table 3, we present analysis of the test-retest reliability for each item, each subscale and the full scale. The table shows the correlation between the same item for the baseline and follow-up survey, which was conducted approximately 3 weeks after baseline. Correlations among the individual items were between .30 and .59. All correlations were significant at p<.0001. Test-retest correlations for the subscales were between .57 and .65. The test-retest correlation for the full scale was .72.

Table 3. Test-Retest reliability.	
ITEM	Correlation between waves 1 and 2
1) You'd better be cautious when dealing with healthcare organizations.	.50 p<.0001
2) Patients have sometimes been deceived or mislead by healthcare organizations.	.40 p<.0001
3) Healthcare organizations often want to know more about your business than they need to know.	.44 p<.0001
4) When healthcare organizations make mistakes, they usually cover it up.	.57 p<.0001
5) The patient's medical needs come before other considerations at healthcare organizations.	.49 p<.0001
6) Healthcare organizations are more concerned about making money than taking care of people.	.59 p<.0001
7) Healthcare organizations put the patient's health first.	.48 p<.0001
8) Patients should always follow the advice given to them at healthcare organizations.	.46 p<.0001
9) I trust that healthcare organizations check their staff's credentials to make sure they are hiring the best people.	.30 p<.0001
10) They know what they are doing at healthcare organizations.	.48 p<.0001
11) I trust that healthcare organizations keep up with the latest medical information.	.38 p<.0001
Agency subscale	.65 p<.0001
Competency subscale	.57 p<.0001
Control subscale	.63 p<.0001
Full scale	.72 p<.0001

Finally, we examined scale validity by testing for an association between the MMI and related measures, the Trust in Physician Scale (TIP) and Global Interpersonal Mistrust-Trust Measure (GIMTM). The TIP is specific to trust, as it relates to patient's trust of his or her physician. The GIMTM measures whether the respondent has a "trusting personality." The MMI was significantly correlated with each measure of trust at .27 and .19 respectively. Both correlations were significant at p<.0001.

List of Publications and Products (We currently have one manuscript under review for publication)

LaVeist, Thomas, Isaac Lydia A., Peterson, Shanni Harris, Jackson John W., Thomas Duane E., "Assessing the validity and reliability of a multidimensional measure of distrust of medical care settings: the Medical Mistrust Index" (under review) *Medical Care*

References

Andersen RM (1968) Behavioral Model of Families' Use of Health Services. Research Series No. 25. Chicago, IL: Center for Health Administration Studies, University of Chicago.

Andersen RM. (1995) "Revisiting the Behavioral Model and access to medical care: does it matter?" Journal of Health and Social Behavior. 36(March):1-10.

Anderson, LA and Dedrick, RF (1990) "Development of the trust in physician scale: A measure to assess interpersonal trust in patient-physician relationships." Psychological Reports 67, 1091-1100

Caterinicchio, RP (1979) "Testing plausible path models of interpersonal trust in patient-physician treatment relationships." Social Science and Medicine 13A,81-99

Couch LL, Adams JM, Jones WH. "The assessment of trust orientation." Journal of Personality Assessment 1996, 67(2),305-323

Flood, Ann (1998) "Risk, Trust, and the HMO: An Editorial." Journal of Health and Social Behavior Sep;39:187-188

Gambetta D (1988) "Can we trust?" in D. Gambetta (Eds) Trust: Making and breaking cooperative relations New York: Basil Blackwell

Hall MA, Dugan E, Zheng B, Mishra AK. Trust in physicians and medical institutions: what is it, can it be measured, and does it matter? Milbank Q 2001;79(4):613-39

Kramer, RM (1999) "Trust and distrust in organizations: Emerging perspectives, enduring questions." Annual Review of Psychology 50,569-598

Mechanic, D, and Schlesinger, M (1996) "The impact of managed care on patients' trust in medical care and their physicians." JAMA 275,1693-1697

Pearson SD, Raeke L. (2000) Patients' trust in physicians: many theories, few measures and little data. Journal of General Internal Medicine 15:509-513

Rotter J (1980) Interpersonal trust, trustworthiness and gullibility." American Psychologist 35:1-7

Sabel CF (1993) "Studied trust: Building new forms of cooperation in a volatile economy." American Psychologist 46:1133-1170

Sleeper S, Wholey DR, Hamer R, Schwartz S, Inoferio V (1998) "Trust me: technical and institutional determinants of health maintenance organizations shifting risk to physicians." Journal of Health and Social Behavior Sep;39(3):189-200

Terrell, F, Terrell SL and Miller, F (1993) "Level of cultural mistrust as a function of educational and occupational expectations among Black students." Adolescence 28(111): 573-578

Thom DH, Bloch DA, Segal ES. and Stanford Trust Study Physician Group (1999A) "An intervention to increase patients' trust in their physicians." Academic Medicine Feb;74(2):195-8.

Thom DH, Ribisl KM, Stewart AL, Luke DA. And The Stanford Trust Study Physicians (1999B) "Further validation and reliability testing of the Trust in Physician Scale." Medical Care. May;37(5):510-7.

Thom DH and Campbell B.(1997) "Patient-physician trust: an exploratory study." Journal of Family Practice Feb;44(2):169-76.

Thorne, SE and Robinson CA (1988) "Reciprocal trust in health care relationships." Journal of Advanced Nursing 13, 782-789

Worchel P (1979) "Trust and distrust" in WG Austin & S Worchel (Eds.) The social psychology of intergroup relations Monterey, CA:Brooks/Cole Publishing

Zucker LG (1986) "Production of trust: Institutional sources of economic structure, 1840-1920" in BM Staw and LL Cummings Research in organizational behavior, vol 8:53-111, Grenwich, CT: JAI Press