TECHNICAL FOOTNOTES

Chapter Two:

44. In organizational behavior research, there is a large, and burgeoning field, which has come to be known as the study of "managerial and organizational cognition," that applies knowledge structuring research to organizational behavior. As a l995 literature review of the field noted (Walsh l995: 281), the topic was the subject of two special issues of the Journal of Management Studies in l989 and l992, a special issue of Organization Science in l994, and of an annual book series called Advances in Managerial Cognition and Organizational Information Processing; in the early l990's as well, the professional association the Academy of Management established an interest group in the area. Noting that people in organizations "(s)omehow…must see their way through what may be a bewildering flow of information to make decisions and solve problems," (Walsh l995: 280), this literature applies theories of knowledge structuring to the behavior of organizations in general and to reactions within organizations to signals from an organization's environment in particular. (see the quotations to this effect in Lindell et al l998: 78).

Much of this literature applies a knowledge structuring perspective to top manager strategy formulation. One of the classic theoretical articles on the topic, Dutton and Jackson's "Categorizing Strategic Issues: Links to Organizational Action" Categorizing Strategic Issues: Links to Organizational Action. Academy of Management Review. l987: 76-90.outlines a set of premises common to many of the articles on the topic. "Attempts to understand the processes through which organizational decision makers learn about the external environment and implement their responses have led many organization researchers to study the link between individuals' cognitive representations of the environment and organizational actions. …According to this perspective, top-level decision makers are bombarded by a continuous stream of ill-defined events and trends. …The meaning of a strategic issue is not inherent in the environmental events or developments. Instead, the organization's internal environment (ideology or structure) has a major effect on the meanings that evolve. …Because organizational actions taken vis a vis strategic issues follow from the meanings attached to these issues, organizations may respond differently to similar environmental events" (Ibid.: 76-77)

Chapter Three

8. In "not broken" were included a small number of respondents coded "no, didn't know any better at the time/wouldn't have thought about it one way or another." Also, the sample of non-supervisory respondents to the In-Person Survey appears to have been significantly biased (due to selection by local office heads) towards supporters of procurement reform. (For a discussion of biases in the In-Person Interviews, see the SamplingIssues.pdf document discussed earlier.) The frequency distribution for non-supervisors for the responses presented here was therefore weighted to correspond to the percentage of those in the Frontline Survey with positive versus skeptical/mixed initial reactions to reinventing government, although the appropriateness of weighting a sample by an opinion rather than by a demographic characteristic is debatable. Responses for supervisors were not weighted.

10. One might suggest that one should use respondents checking the second or third box for classifying earlier recruits, since the wording of these response alternatives corresponds quite closely to what it means to have been an earlier recruit. The problem is that it is possible that some who were critics of reform as of the time of the survey (particularly those giving the response, "I have mixed feelings about acquisition reform. There are some things I like and some I don't") might initially have been early recruits but become disillusioned since then. The only response categories that would seem to be inconsistent with having been an early recruit would be categories 4 and 5, where respondents reported themselves as having "for a long time" been "skeptical about whether this was the right way to go." This group constituted only 6.5% of the sample. Respondents who gave this response and also disagreed they initially tried reform only to please their bosswere providing inconsistent answers (this constituted 1.2% of the sample), so a decision had to be made about whether to include them as early recruits (based on their response to initial reaction) or not. Here, and in the models to be presented below, the decision was made to include them as early

recruits; the results otherwise would have been close to identical. In the case of the regression models to be presented later, I also ran these models excluding the overall attitude=4,5 respondents who had values of initial response=4,5. Results were generally more or less identical to the models to be presented in that table; there was one situation where there was a difference large enough to be worth noting, which will be discussed in a footnote.

[14] As critics of Berlin (for example, MacCallum l967) have noted, there is some artificiality in this distinction because every liberty in fact involves both "freedom from" and "freedom to." So the desire for autonomy for the sake of burden reduction may be seen as freedom from rules and hierarchy to avoid doing things one would prefer not to do. However, focusing on "freedom from" or "freedom to" suggests which is most central, problematical, or difficult to achieve.

20. A motivated desire for autonomy generally involves a substantive desire to choose or to produce some substantive better outcome through one's actions. However, as Dworkin (1988: esp. Ch. 2) forcefully argues, it may involve only valuing the ability to make a "meta"-choice about life plan, whatever that life plan is, rather than specifying that the plan must be rich in opportunities for choice. (Ibid.: 31) "(T)he person who decides to do what his community, or guru, or comrades tells him to do" can still value autonomy in this sense, though he or she doesn't seek to make lots of individual choices himself or herself (Ibid.: l2) "It is his decision, backed by reasons, arrived at freely. …He is leading just the kind of life he thinks is worth living." (Ibid.: 23) Dworkin's argument is analogous to the contention in Nozick's Anarchy, State, and Utopia (l974: Ch. 10, esp. 320-23) that libertarians might, in organizing their personal lives, choose to live in communities placing significant restrictions on their freedom to behave as they choose in specific situations.

Applied to individuals making decisions for themselves, the argument for autonomy based on a desire to achieve better results (in one's own life) features prominently in philosophical discussions, particularly those criticizing paternalism. (Mill 1859: 81-82) It is commonly argued that, empirically, people tend to make better choices for themselves than others would make for them, since they are more knowledgeable about their own preferences and have a greater incentive to make appropriate choices for themselves than does a third party, who would be presumed generally to be less concerned about their welfare than they themselves are.

If an individual is rewarded within the organization for making good decisions for the organization, then the anti-paternalist reason to desire individual autonomy would apply to job autonomy – one would desire job autonomy because it increases the chances one will make decisions that produce better organizational results and hence individual rewards. Furthermore, since anyone (more or less) can follow a rule, buck a decision, or do a highly specialized job, a bureaucratized system doesn't allow people to distinguish themselves from others as much as one based more on judgment.

.A person seeking job autonomy to succeed better in his or her organization might well (indeed, might be expected to) use autonomy to produce better results from the perspective of the goals currently valued in the organization – perhaps, for example, to come up with smart ways to police program officials or contractors better. An interest in job autonomy based on the desire to be successful at one's job would thus by no means necessarily involve support for the procurement reformers' agenda.

21. Converse coined the phrase "level of conceptualization" to refer to a sliding scale expressing such variations in the degree of remoteness, generality, and abstraction (Ibid.: 2l5) In The American Voter (Campbell et al 1976: 250), which was part of the same research program, the authors presented evidence differences in education were a major determinant of variation in individuals' levels of conceptualization. In their data, 5% of those with grade-school, 10% with high school, and 32% with college education were classified as displaying ideological or near-ideological levels of conceptualization.

Re-examination of the question of levels of conceptualization by scholars in the l970's (Nie et al, 1979), writing during an era of greater ideological divisions than the l950's, called attention to the point that levels of conceptualization might vary over time, as ideological issues became more, or less, salient in political debate. They argued that Converse's results were partly explained by the non-ideological tenor of political debate during the Eisenhower era of the l950's. However, although the gap between the proportion of those using more abstract concepts among those with at least some college and those with less than a high school degree narrowed between l956 and l972, it was significant in both years (ibid.; 121).

23. Inglehart distinguishes between "postmaterialists" and "materialists" using a number of survey questions, but his most basic ones ask respondents which of four goals was most desirable and second-most desirable to attain (1971: 994): (l) maintain order in the nation, (2) give people more say in important government decisions, (3) fight rising prices, (4) protect freedom of speech. Respondents listing the second and fourth goals as most important are categorized as "postmaterialists," those listing the first and third as "materialists," with others as mixed cases. Inglehart argues that the postmaterialist values he measures involve self-expression and the ability to choose, both of them versions of autonomy in the political sphere that are closely related to the opportunity to exercise judgment/choose on the job.

In later surveys, in order to take account of people growing up rich in poorer societies and vice versa, he looked at the respondent's report of his or her family's socio-economic status when the respondent was growing up (parents' education and occupation at the time the respondent was growing up). He found similar effects with this specification. It should also be noted that Inglehart disagrees with the view that the relationship between increased wealth and postmaterialism is fully or even partly mediated by education – that is, that as societies grow more wealthy, young people are provided more education, and that it is greater education that (for reasons discussed above) increases postmaterialism, alone or in addition to affluence. (Abramson and Inglehart (l995: 76) Frustratingly, however, Inglehart's work itself has never taken the simple step of undertaking even ordinary least-squares regression analysis to test his view. A zero effect of education controlling for childhood economic status in such a regression equation would allow rejection of the hypothesis that education has an independent impact on postmaterialism; a zero effect of childhood economic status controlling for education would lead one to reject the hypothesis that something about childhood economic status itself has an independent influence on postmaterialism. Two existing studies (Dalton l977; Duch and Taylor l993) perform such a regression analysis. Dalton (l977: 466-68) concludes that both "the economic environment surrounding the cohorts at age l0" and education levels have an independent influence explaining postmaterialism, with childhood economic environment somewhat more important, while Duch and Taylor (l993: 763) conclude education is important but not childhood economic environment, which in their regression analysis actually has a sign opposite from the one Inglehart predicts. The value of these analyses is limited by the lack of individual-level data. Both studies use of group-level averages for each of the variables, whereby an entire age cohort in these surveys becomes the basis for each variable (e.g. postmaterialism level for an age cohort, average educational level for an age cohort). This tends to exaggerate correlations and sometimes can be misleading.

25. One of the rare examples of – passing – references to this idea in the literature is in Kanter (l983: 24, 56), who states job autonomy is "better suited to the demands of our times for more responsiveness to the needs of educated, rights-conscious employees," and that "education creates pressure for autonomy, flexibility, and freedom." Another is Lawler's observation (l992: 41) that "most employees prefer to work in an environment where they are involved in the decision-making process; have challenging, interesting work; and are expected to be more than a pair of hands. This is particularly true of professionals such as engineers and scientists. …" Looking back into the past, when blue-collar workers were even less-educated, Lawler (l992: 26) notes that Taylor's bureaucratic approach to organizing the workplace was "ready-made for an American work force that consisted of poorly educated immigrants." Also, scholarly accounts of professions --occupations consisting of highly educated people – often emphasize autonomy as a central feature defining these jobs. (e.g. Freidson 1970: Ch. 2)

26.In examining personality trait variables (also called in social psychology "individual differences" or "dispositions"), we should remember that social psychologists generally tend to be skeptical of the impact on such variables on behavior, arguing that situational factors are considerably more important than personality differences. For a strong statement of this view, see Ross and Nisbett l99l. For critiques of this view, see, for example, Bem and Allen (l974), Funder and Ozar (l983), and Jouse <u>et al</u> (l996). The dependent variable here is an <u>attitude</u>, however, to which these skeptical strictures might not apply in the same way. For a study that finds impacts of dispositions on job satisfaction attitudes, see Staw <u>et al</u> (1986).

34. To a limited extent, the government provides assistance to employees seeking higher educational degrees (more commonly to employees in higher-level positions), and employees may also seek these degrees on their own. Since the value on the education variable is as of the time of the survey, it is likely some respondents received additional degrees after being promoted to their current positions. To the extent this occurred, this biases the coefficient for job level upwards. (There would be an omitted variable bias: the omitted variable is "likelihood of receiving an education after job promotion" that is positively related both to job level, the dependent variable, and education level, the predictor variable.)

37. Ingelhart (1990: 88-90) discusses lifecycle effects as an alternative explanation for his results and argues evidence from his surveys does not support that alternative explanation.
:
42. The missing in action status of these responses is even more dramatic if current-attitude bias exists in the In-Person Interviews, which might have caused some respondents to remember these as criticisms of the traditional system even if they hadn't experienced them so at the time, since these became prominent themes after reform began. This, incidentally, may be seen as additional evidence that memory distortion is not an overwhelming problem with the recall data in these interviews. These problems were mentioned, though occasioinally by frontline employees, in other parts of the interviews.

Chapter Four

1. Interestingly, the antibiotic study reached many similar conclusions as the earlier seed studies, but the sociologists who did the former research were unaware of the research, by agricultural economists, on diffusion of the use of new seeds (Rogers 1995: 68)

2. There are analogies and disanalogies between phenomena discussed in diffusion of innovation literature and earlier support for procurement reform. This literature generally examines individuals who, although part of informal social systems, are not parts of a formal organization, and they are generally making individual decisions about whether to adopt a new technology or practice. No decision for a collectivity is needed to get change underway. (An individual doctor can successfully use a new antibiotic even if no other doctors do.) The diffusion of innovation literature also takes a behavioral variable – use of the innovation – as the dependent variable, while in this discussion, we take attitude towards the innovation as the dependent variable. However, if these differences are combined, the circumstances studied in the diffusion of innovation literature actually start seeming more analogous to issues involving earlier support for procurement reform. Although getting behavioral change underway in a formal organization may significantly be a collective process beyond the control of an individual, attitudes may easily be adopted individually, as innovations can. Another possible disanalogy between this literature and the initiation of procurement reform is that in the case of reform, there was a change vanguard, who had "adopted" the change even before it become available. However, many early adopters of technologies showcased in the diffusion literature may have, in general terms, adopted the generalized "change" of being oriented towards new ways of farming or new ways to treat patients, like the change vanguard in the procurement system, which then prepared them to be early adopters of the specific innovations of hybrid corn seed or tetracycline. Indeed, diffusion researchers argue that early adopters are pro-innovation in general.

14. Since standardized coefficients for interaction terms are difficult to interpret, the table lists non-standardized as well as standardized coefficients, so we can interpret interaction effects more readily. For these and all subsequent models, both standardized and non-standardized coefficients will generally be given to the nearest two digits; however, for some variables taking many values, where non-standardized coefficient for a one-unit move is very small, non-standardized coefficients are presented to the nearest four digits.

18. In calculating the mean values on seeing reform as being about reducing burdens on contracting people, I excluded respondents disagreeing with that statement but agreeing with one to the effect that their first reaction was that reform was mostly about reducing burdens on <u>industry</u>. I wanted to look only at respondents who initially saw reform as involving either making their jobs easier or as involving some other benefit for the government. If I hadn't excluded respondents agreeing with the statement about

initially seeing reform as involving reducing burdens for industry (these would largely be people supporting the ideology of the traditional system), I would have mixed respondents who disagreed with the statement because they thought reform meant buying faster or better value (which is the contrast I wanted to examine) and those who disagreed because they thought reform meant reducing burdens for industry.

The number of responses in this table is rather small, because l) 22% of respondents didn't answer one of the two questions used to develop this category (not surprisingly – in fact, reassuringly, a larger percentage of respondents left retrospective recall questions blank than most other questions; 2) the only respondents included were those agreeing and disagreeing (respectively) with the two statements used to develop the category, leaving out the many people who answered "3" ("mixed feelings") to at least one question.

21. The result from the In-Person Interviews presented in the footnote compares supervisor and non-supervisor respondents who both stated they would have said in l992 the system was broken (or had significant problems) and also were coded as having had a favorable initial reaction to announcement of the "reinventing government" initiative. Of 27 non-supervisors in that category, two gave the "best value" response as a reason they thought the traditional system was broken. Of eight supervisors who did so, three gave that response. The usual calculation of a p-value based on a normal approximation was not appropriate here because of the small number of respondents, so instead a t distribution with 33 degrees of freedom (35 minus 2) was used to perform the two-sided hypothesis test.

25. These results provide evidence that both idealism exerted an impact via a respondent's general ideology and not just via "diffusion of innovation"-type impacts on willingness to support lost causes or ability to conceive of an alternate world that doesn't already exist. The reason is that the impact of ideology provides a plausible story of why there should be an interaction effect between them, while for the lost cause path, there is no plausible story why support of lost causes and job level should have a multiplicative effect, though such an interaction might exist in terms of the ability to conceive of an alternate world. This doesn't mean "diffusion of innovation" paths exerted no influence, only that the path ideology did.

28. We shall see below that a far larger percentage of supervisors than non-supervisors were earlier supporters of reform, reflecting higher values for other variables in the model predict earlier support.

29. It is possible that answers to timeliness wholly or partly might reflect a source of personal job burden/stress rather than support for a TQM customer service ideology, as suggested above: a respondent may have perceived timeliness as important because slow service subjected the person to criticism from program people, independent of how much the respondent valued customer service. To test for this alternative explanation, the model tested for an interaction effect between this variable and a question about whether people found it stressful to say no to program customers, a variable that was not significant as a main effect in any model. If dissatisfaction with slow service was driving earlier support for reform because it created personal burden/stress, one would expect to see a stronger relationship between dissatisfaction with timeliness and earlier support for reform among those finding having to tell program people about delays more personally stressful than among those who didn't. The interaction was not significant.

31. Longevity might have exerted an impact only via attitudes on other variables, such as other resistance to change variables or support for autonomy. However, there was no first-order correlation in either model between the dependent variable and job tenure (it was .00 and .0l), so the lack of effect was not simply a result of controlling for other variables.

Another possibility is that other impacts of longevity balance out a resistance to change effect. People with more seniority may experience a freedom to engage on behalf of their beliefs that lower-level employees may not feel, a phenomenon that may also help explain the over-representation of supervisors in the change vanguard, to be discussed below. One may speculate that those who've been around an organization for long enough may develop both a sense of independence, perhaps a sort of "serenity," coming from being relatively unlikely to be promoted beyond whatever position they have already attained. Thus, they became more willing than those with less independence to join a change movement.

The view there are factors counterbalancing influence of longevity on resistance to change is supported by the only other study I was able to find examining this question empirically (Trumbo l96l: 341), which examined general attitudes towards job-related changes among a supervisory and non-supervisory employees at an insurance firm. It found attitudes to change did not vary by respondent age, though the study did not discuss why that might be.

35. Another possibility is that results for the workgroup were mistaken and that in fact members of the change vanguard were not succeeding influencing people to become early recruits. Values for this variable account for people currently in the workgroup who had not been there in l993, but they don't account for people who left after then. To the extent people leaving were disproportionately not members of the change vanguard, the value for this variable is too high. In addition, assuming that differential sample attrition was the same across workgroups, workgroups with a lower percentage of people actually in the change vanguard at the time will have lost more reform skeptics to sample attrition because there had been a higher percentage of such skeptics in the workgroup (if 80% of the workgroup was not in the change vanguard and there was a l0% differential sample attrition, that workgroup would have lost eight people to sample attrition, while if 70% were not in the change vanguard, the workgroup would have lost only seven). This means that the bias in these data is that the lower the reported value of this variable, the greater the extent to which he reported value overestimates the real value. This means that, if there is a positive relationship between this variable and earlier support for reform, the actual slope of the relationship is flatter than the reported slope (because lower values for x are more exaggerated than higher values for x), meaning that any reported relationships between this variable and earlier support for reform exaggerate the real relationship. (I am grateful to my colleague Suzanne Cooper for pointing out to me that changes in workgroup composition since l993 constitute a potential source of bias in this variable.) At the same time, the use of buying office-level data for those respondents for whom data was not available at the division level presumably biases any results in a conservative direction because, especially in the larger organizations, the buying office as a whole probably consists of many workgroups. Indeed, taking the workgroup at the division level may produce similarly conservative results, because in a number of these organizations a natural workgroup would be smaller than the division level. In all, it would seem unlikely that the biases created would be so strong as to render a coefficient with a p-value of .00l statistically insignificant.

A word is also in order about possible biases in coefficients for most-respected coworker attitude. For respondents reporting their most respected coworker was an earlier supporter of reform, this was a retrospective recall question, while for other respondents reporting a more recent attitude on the part of the most-respected coworker, it wasn't. Current-attitude bias suggests that respondents recalling their most-respected coworker was an earlier supporter of reform may have been exaggerating that person's degree of earlier support, meaning that the actual mean value for the extent of early respected-coworker support for reform among initial reform supporters was higher (less pro-reform) than the reported value, making the slope of the actual regression line steeper and reported coefficients too conservative. As with keeping up with top leaders, there might also be an endogeneity problem with most-respected coworker attitude.

48. The most-respected coworker attitude percentages count outside these two categories the 29% of respondents stating they didn't know the opinion of the person whose attitude they respected the most, so, of those who stated they knew the opinion of their most-respected coworker, 57% characterized that person in one of these two categories. It is possible some coworkers became the person the respondent respected most because of their support for reform, in which case percentages here would be exaggerated.

49. Because of issues, discussed earlier, about the status of post-l993 "team leaders," these respondents were omitted from these mean values. Note presence of the so-called "Lake Wobegon" effect, confirmed in numerous studies where people are asked about driving and other behaviors, where the average person believes he or she gets more requests than average.

62. Just to cite one example, though there is a significant literature on this phenomenon, a study of participants in the l988 party presidential nominating processes showed that while 38% of Democratic voters described themselves as liberal, 68% of delegates to state-level conventions did so – and the corresponding figure for Republicans describing themselves as conservative was 65% and 9l%. (McCann l996: 79-80) Party activists who take the trouble to attend presidential primary caucuses are likewise more ideologically polarized than party voters; in one study, 59%, 58%, and 77% of Republican caucus attendees in three states described themselves as "conservative" or "extremely conservative," compared with 39% of all Republican voters; figures for liberalism among Democrats were analogous. (Mayer l996: 134-35) A study of party presidential nominating convention delegates as early as l972 showed that over 80% had completed at least some college. (McCann l996: 75-77) In three states having presidential primary caucuses, l8%, 31%, and 27% of Republican caucus participants had graduate school degrees, compared with 6% of general election Republican voters, and again figures were similar for Democrats. (Mayer l996: 131-32)

63. Writings such as those of Lenin and Walzer present two ideal types: the leader who is well-educated, economically comfortable, and politically visionary, and the follower who is less-educated, poorer, and politically focused on concrete improvements in his or her personal life. These writings contain nothing resembling a well-specified causal model of relationships between being a revolutionary leader or follower on the one hand and other variables such as education, social position, and interest in abstract versus immediate concerns on the other. Probably what underlies these arguments are causal patterns growing out of the sociology of societies at low average income levels, where the poor were on the verge of starvation and where education was not widely dispersed. The poor had very low levels of education and thus an inability to think abstractly, while they were close enough to the margin of existence so the desire for short-term amelioration assailed them with some urgency. The wealthy were educated and thus had an ability to conceptualize abstractly, while not needing to worry about starvation, so those among them who were discontented gravitated towards abstract visions of a new society.

Fast-forward to the modern world. Obviously, no procurement employee in contemporary America is near the margin of physical existence. Overall education levels are much higher. While the procurement change vanguard seems to have shared the same discontents with early recruits but had an additional source of discontent as well, for many revolutionaries, ideological discontent substituted for more everyday discontents. Ideology drove revolutionaries to high levels of activism (as it would appear to drive party activists in contemporary America), while there is no evidence that the ideological views of members of the change vanguard preoccupied them; indeed, all we can say is that there is evidence that members of the change vanguard were more likely to embrace a job requiring original work, less likely to see reform as involving reduction of job burdens only, and more supportive of the reformers' better value from contracting agenda. Correspondingly, members of the change vanguard did not, unlike revolutionary leaders, actively organize an oppositional movement on their own.

Chapter Five

14. In revolutionary situations in oppressive political regimes, the shock may in significant measure involve something as elemental as elimination of government repression or military collapse, as in Russia in l9l7. But opportunities the top of the system suddenly provided the discontented at the bottom have often not been limited to disappearance of repressive force. When France's many wars had driven the government to the brink of bankruptcy and lenders would no longer provide funds the Crown needed, the King called the Estates-General into session for the first time in hundreds of years, an event that was to set the French Revolution in motion. When he called the Estates-General, the King invited local communities to develop and present to the government lists of grievances. In communities throughout France, unprecedented meetings, often at local churches, occurred where people were authorized express their discontent. Many members of the nation's elite publicly expressed support for change. Soon, the King began referring to himself as a Citizen-King. (Schama l989: Ch. 8) The opportunity thus included authorization to express dissent, along with support for change coming from many traditional leaders.

15. The natural selection metaphor is at the center of a prominent approach to organization theory called organizational ecology, which argues the universe of organizations looks different over time as organizational environments change, not because any given organization adapts to a changing environment, but because exogenous changes in environments favor those already having characteristics suitable to the new environment. However, the unit of analysis in this theory is the organization as a whole, rather than subcultures whose importance might rise or fall due to selection pressures. The use of natural selection theory to explain successful initiation of organizational change actually is an argument against the conclusion in organizational ecology theory that existing organizations seldom can successfully change.

We should also note the discussion in Van de ven et al (1999: 28-30) on the role of "shocks" in triggering organizational change. Summarizing from their empirical examples, they write: "Shocks were important in each case because they allowed the champions of an innovative idea to gain currency with various potential stakeholders within the organization. Even though the entrepreneurs or champions were often convinced about the potential of their ideas, the rest of the organization did not necessarily share this 'insight.' In the typical scenario, the champions rarely controlled the resources required to develop their insight or ideas. In most of the cases studied, an opportunistic champion could not move the innovation forward. The idiosyncratic vision or insight of the champions was not widely shared in the rest of the organization. Indeed, potential stakeholders had to be convinced to support an idea." (Ibid.: 29) As the quote suggests, the account is different from the one being presented here. Although it emphasizes the role an exogenous shock can play in helping people on the front lines of the system, the mechanism by which this proceeds is predominantly to make it easier for innovation advocates to "sell" their idea to people higher up in the organization.

The activating the discontented theory presented here (like the theory of evolution, it may be noted) is a process theory. (Mohr l982: Ch. 2; Poole et al 2000) By this is meant that it "tells a…story about how something comes about." It includes various elements in a temporal order that "(combine) with one another in such a way as to yield the outcome." (Mohr 1982: 44-45, emphasis in original) The individual "(i)ngredients alone do not convey a sense of explanation. …There must also be some instruction for mixing them – a recipe. Recipes generally mandate activities that occur over time and in a prescribed order." (Ibid.: 60) So, for example, in the theory of evolution, the first ingredient is the presence of a certain trait in the population. The presence of that trait, however, will not determine whether that trait becomes widespread in the population (the dependent variable) unless there is a subsequent change in the environment that favors individuals possessing that trait. In a process theory, the linking of various elements is more important in explaining the dependent variable than in theories where there are a number of unconnected predictor variables and the dependent variable is explained by changes in the value of the predictor variables.

Chapter Six:

3. It may be noted that, by contrast, the multiple-regression oriented social science literature finds relatively small coefficients for the role of variation in leadership support in explaining relative success in organizational innovation. (Damenpour 199: 576) However, in these studies, "leader support" is typically measured through a very simple single-question variable such as the orientation of organizations' leaders towards change in general; these studies do not measure leader behaviors.

17. McGregor contrasted "Theory Y" to Theory X. (Ibid.: 47-48) According to Theory Y, "The expenditure of physical and mental effort in work is as natural as play or rest. …(W)ork may be a source of satisfaction (and will be voluntarily performed)." Under proper conditions, "(t)he average human being learns…not only to accept but to seek responsibility." Employees committed to the goals of the organization will use autonomy to try to do a better job. Hackman and Lawler (l97l: 262-63) write about "motivating jobs," that achieve "a congruence between individual need satisfaction and organizational goal achievement." In such jobs, "(t)he harder and better an individual works on such a job, the more opportunities he will have to experience higher order need satisfaction."

Furthermore, if employees want the opportunity to use their minds on the job and more such opportunities are provided them, this may generate greater commitment to the organization and more discretionary work effort. One meta-analysis (Spector l986: 1009-ll) of empirical studies relating employee self-reports of the extent of job autonomy to self-reports of employee commitment to the organization for which they worked and to their degree of psychological involvement with the job, as well as to job performance (measured by supervisor ratings of the employee) found positive correlations (.23, .l7, and .20 respectively). In the Hackman and Lawler study (l97l: 279), for respondents with great "higher-order" need strength, the correlation between the degree of job autonomy and the quality of the respondent's job performance (rated by a third party) was .l7, and with the quantity of work it was .09. For those with low strength for such needs, the correlations were -.05 and -.09; those seeking burden reduction from job autonomy seem to have reacted to gaining autonomy by working less. One must be cautious about correlations among self-reports of various job characteristics where one set of self-reports is assumed to be predictor variables and another self-report is assumed to be the dependent variables; all the self-reports may reflect an underlying dimension of overall attitude to the job. (Tosi et al 1976: 300; Salancik and Pfeffer 1978; James and Jones 1980). Reverse causation is also possible: job satisfaction makes one believe one is experiencing these various positive job attributes. A second form of reverse causation is also possible: employers may be willing to give employees with greater commitment to the organization a greater degree of autonomy.

19.The literature on innovation refers to the "re-invention" of an innovation as it is implemented (Rogers l995: 176-79). This involves either customizing the innovation to fit local circumstances or picking for implementation a few features from the many the innovation included. A number of studies of public school innovations have noted that the greater the re-invention, the more likely the innovation was to be sustained rather than discontinued. (Berman l977) As Van de Ven et al (1999: 57) note in comparing, as change initiation strategies, in-depth pilots of an entire innovation with broad testing of a shallower version of the innovation: "Fewer hurdles and resistances to change are encountered when a few, presumably easy, components of an innovation are implemented across the board to a few, presumably supportive, stakeholders, than when all, easy and hard, components of a program are implemented in depth with all partisan stakeholders involved."

Chapter Seven:

9, Criticisms such as these have been made, for example, of the literature, based on surveys given at one point in time, showing a positive connection between variables such as on-the-job participation or closeness of supervision and job satisfaction. These concerns are partly, though not completely, behind the worry about what is called "common method variance," (Podsakoff and Organ l986) exaggerated correlations of predictor and dependent variables gathered from the same self-reports that may reflect some underlying dimension -- though not necessarily perceptual confirmation, which is more of a causal phenomenon.

The theory of "perceptual confirmation" in psychology would suggest that, at least to some extent, such a phenomenon in fact does tend to occur. An example of an experiment showing operation of perceptual confirmation specifically in the context of job design is Staw 1975. In the experiment, two sets of groups performed a team exercise. One set was randomly informed afterwards that they had "done quite well" at the exercise, while the other was informed they had "not done too well." (In fact, there were no overall performance differences.) Asked afterwards to characterize various aspects of the team's performance, such as team cohesiveness, communication, and openness to change, the teams informed they had done well reported above-average scores on these team attributes than those informed they had done poorly. Team members apparently had the idea that teams high on cohesiveness, communication, etc. would perform better. Believing they had done well, they then engaged in perceptual confirmation, perceiving team characteristics corresponding to their pre-conceptions about successful group experience.

A better way to have dealt with endogeneity would have been to use two-staged least-squares regression. However, there were no plausible instrumental variables that could be used to develop such a model. As noted in Chapter Five, this creates the risk that coefficients reported here are somewhat distorted.

21. Actually, Merton described self-fulfilling prophecies as involving only false predictions, a limitation that has been, rightly, rejected by subsequent researchers. (See, for example, Jones l977: 4.)

23. Some methodological problems in the initial Rosenthal and Jacobson study caused it initially to be quite controversial. Subsequent replications of the effect in other studies have increased confidence in the findings. (Jones 1977: 107)

24. As these examples indicate, the self-fulfilling prophecy may operate both interpersonally and intrapersonally. Most of the examples in Merton's essay involve self-fulfilling prophecies in interpersonal contexts. This was consistent with Merton's interest in situations where one group's ethnic prejudices against another encouraged behaviors among the victim group that fulfilled the prejudice. The intrapersonal self-fulfilling prophecy forms the basis of the classic children's story, The Little Engine That Could, where the little engine repeated the mantra, "I think I can, I think I can," to help summon forth the effort to allow it successfully to perform the dauntingly difficult task of making it up the mountain.

26. A general discussion of the operation of self-fulfilling prophecies in organizations, including a discussion of the limited then-existing empirical work on the topic, is Eden (1984). The dependent variable in Eden's discussion is always subordinate performance based on supervisor performance expectation; none of the discussion deals with a person's own expectations.

30. Comer and Laird (1975) suggest that the source of the tension is that people seek to make sense of the events of their lives. The drive to reduce cognitive dissonance occurs, in this view, because people seek to make sense of their behavior, giving a reason for it (i.e. a corresponding attitude), rather than seeing it as arbitrary. A related suggestion (Staw 1980: 56-57) is that the gap between one's behavior and attitudes damages one's self-concept and hence induces dissonance.

31. Generally, cognitive dissonance theory argues, somewhat unhelpfully, that people will change those cognitions that are easier to change. Attitudes are typically easier to change than behaviors because they are more ambiguous and easier to deny, since they are less objectively observable. Behaviors one has undertaken more than attitudes one may also be more closely tied with a person's self-image (Kiesler 1971: 43) and thus more resistant to revision.

35. One further element of the theory should be noted. Dissonance theory argues that, for dissonance to be triggered, the individual must feel a connection to the behavior they've undertaken, typically through the individual's perception they have chosen to undertake the behavior. (Sherwood et al l972: 83-84; Wicklund and Brehm 1976: 56) If an individual feels he or she undertook the behavior because forced or bribed, without personal volition, the theory argues, the behavior won't arouse dissonance.

Presumably, a person finding himself committed to doing something contrary to his attitudes for a large reward or from coercion can deny responsibility for his behavior. He can externalize the reason he is doing what he is doing; that is, he can say to himself: "I don't really believe in this, but I really had no choice because I cannot afford to refuse such a large reward; therefore, I am justified in doing this even though I believe it is not what I should be doing." On the other hand, a person who receives a minimal reward or very slight coercion cannot justify his attitude-discrepant actions so easily. He is more likely to conclude: "I got myself into this situation, and because I don't normally do things in which I don't believe, there must really be something to the position I am advocating." (Sherwood et al l972: 84)

This part of the theory should raise a red flag: if dissonance-arousal requires choice, why would it ever occur in the case of attitude-discrepant behaviors? If they don't choose to do so, and if choice is required for dissonance-arousal, then how can this theory apply to attitude change brought about by attitude-discrepant behavior? The answer comes from a closer examination of experiments involving such behavior. In virtually all of them, the main choice subjects made was to agree to participate in the experiment in the first place, which presumably included a voluntary commitment to do what the experimenter asked. People then accepted participating in the painful initiation or even in caterpillar eating. This suggests pressures for attitude change in the presence of attitude-discrepant behavior exert force even absent whole-hearted, fully voluntary initial behavioral commitment.

Often, though surely not always, initial skeptics/fencesitters would have felt as strong (or as modest) an element of voluntary choice in decisions to display reform-oriented behaviors, despite initial skepticism, as did subjects in dissonance experiments. Like the subjects, their main voluntary choice would have been a meta-choice – the decision to join (and stick with) the organization in the first place, just as had agreed to participate in the experiment. Furthermore, as Herbert Kelman has argued (1974: 321-24), a person not fully convinced a certain new attitude is correct, but intrigued by the possibility it might be, may voluntarily choose to try out a new behavior, even while skeptical, to "see how it works" – and then become non-consciously subject to the attitude-changing impacts of cognitive dissonance (or of other mechanisms producing attitudes from behavior).

39. These two mechanisms would appear to operate via a phenomenon the psychologist Darryl Bem (1972: 2) has dubbed "self-attribution." (In social psychology, "attribution" refers to judgments about how we explain people's behavior (Myers l999: 80). "Self-attributions" are one's own attributions about the causes of one's behavior.) The argument is that individuals come to "know" their own attitudes partially by inferring them from observations of their own overt behavior. "(T)o the extent that internal cues are weak, ambiguous, or uninterpretable, the individual is functionally in the same position as an outside observer, an observer who must necessarily rely upon those same external cues to infer the individual's inner states." In their original "foot in the door" article, Freedman and Fraser (l966: 20l, emphasis added) explained the operation of the mechanism as follows:

> Once he has agreed to a request, his attitude may change. <u>He may become, in his own eyes, the kind of person who does this sort of thing</u>, who agrees to requests made by strangers, who takes action on things he believes in, who cooperates with good causes.[i]

Or, to use Karl Weick's (1979: 5) renowned aphorism, "How can I know what I think until I see what I say?"

These mechanisms are often presented in the context of behavioral change – one becomes willing to undertake "bigger" behaviors after having done "smaller" ones. However, according to the self-attribution explanation for these phenomena, behavioral change occurs because the first behavior generates an attitude that then in turn generates further behaviors, so the mechanisms operate by generating attitude change. "Self-attribution" theory was originally developed as an alternative explanation to cognitive dissonance theory for some of the same experimental findings, and a number of experiments test the two theories against each other as rival hypotheses. This dispute need not concern us here; I will emphasize instead the different kinds of people and situations where the different theories are likely to be most applicable.

Wicklund and Brehm (l976: 280) see self-attribution theory as an account of attitude <u>development</u> more than attitude <u>change</u>. Additionally, foot in the door is less dependent on voluntary choice than cognitive dissonance, so it might affect skeptics/fencesitters not seeing the decision to participate in reform-oriented behaviors as voluntary.
In a later experiment (Snyder and Cunningham l975), the self-attribution explanation for the foot in the door effect was supported over an alternate explanation, which was that subjects who had already received a small request would find the large request relatively less onerous than those who had never previously received any request. In this second experiment, subjects who had previously turned down a larger request were less likely to comply with a moderate-sized request than were those who had previously accepted a smaller request and now asked to comply with the identical moderate-sized request – a result opposite from the one predicted by the "relatively less onerous" hypothesis.

51. I observed (Kelman l981) an analogous phenomenon in studying popular support for social democracy, a highly egalitarian movement, in Sweden. Part of the source for support for egalitarianism in Swedish politics came, I argued, from an earlier Swedish tradition of deference to government authority, which got transferred to the Social Democratic government after it served decades in power. In effect, the inegalitarian tradition of respect for authority produced obedience to the government's order that people support egalitarianism!

53. I am grateful to my colleague Nancy Katz for helping me with this distinction and for suggesting the examples given here. In the sample, the correlation between self-confidence and the variable measuring

external locus of control was -.l6, indicating there was a relationship between high self-confidence and an internal locus of control, but the relationship was not overwhelming.

Chapter Eight

4. Besides these “beginning” and “end” variables, the Frontline Survey includes no other “in-between” measures of experiences and attitudes, for obvious reasons of measurement error in the recall of a series of unanchored events in the past. With this limited number of datapoints, the coefficient for the self-fulfilling prophecy in the model using empowerment as the dependent variable (which will be the one used to measure determinants of overall successful experience with reform) is likely to be too low and those for other positive feedback mechanisms too high. The reason is that these other positive feedback mechanisms (say, for example, self-confidence or behavioral facilitation) have two effects, directly, to increase the quality of each “intermediate” experience with reform and, indirectly, thereby to improve expectations for the experience with reform immediately following it, which in turn, through the self-fulfilling prophecy, increases the quality of that next experience. The coefficients for these other positive feedback variables will measure both their direct impact on successful experience and their indirect effect via operation of the self-fulfilling prophecy. The indirect effects are not fully controlled for by the variable (initial experience) used in the empowerment model to measure the effects of the self-fulfilling prophecy, because we only have that early datapoint measuring the presence of expectations that produce the self-fulfilling prophecy, not later measures.

Also, various of the positive feedback effects discussed earlier were theorized differ in strength at different stages of the change process. One would expect that effects due to mere initiation would tail off over time, but it is an empirical question when this occurs, and also whether these effects are at their greatest at the very beginning of a change process or somewhat later on. Effects due to the mere passage of time might well also differ in strength at different times, and might also tail off. With only two datapoints, no conclusions may be drawn about how these different effects work at different stages of a change process.

It should be noted that the same cautions applied to interpretation of data in the earlier support models discussed earlier apply to interpretation of coefficients for initial experience as well, though with a bit less force because the sample included two additional years of respondents.

Measures such as mean workgroup attitude, most-respected coworker attitude, supervisor attitude, and local office head attitude are one-time measures based on current attitudes of others in the respondent’s environment, which in some cases (such as mean workgroup attitude and supervisor attitude) were likely more positive towards reform than during the period of the change process as a whole. However, if these rose over time equally across workgroups and supervisors, one would not expect these changes to affect coefficients in the model.

Current-attitude bias, as noted earlier, would produce a situation where the coefficient for initial experience, compared with other coefficients in the model, would be too high. (The same would be the case for initial experience as a predictor variable in the empowerment model.)

I also ran the initial experience model excluding members of the change vanguard, on the grounds that their positive reactions to their first experience would be less interesting to explain. However, the results in the two versions of the model were essentially identical, so results for the full sample (excluding those who had started working in their organizations after l994) are given, to permit greater ability to compare with the models explaining earlier support for reform presented earlier.

6. To make this calculation, I excluded the control variables. Since making these exclusions, which was necessary in order to exclude control variables from the calculation, means that all covariance got assigned to the variables remaining in the models, these R-squared numbers are too high.

9. The overall effect was insignificant However, a quadratic term for working late was significant, thus showing nonlinear effects. However, the form of the nonlinear effect was a U-shaped curve, where successful experience was highest for those who strongly agreed that they worked late and who strongly disagreed, with poorest experience for those in the middle. So working late promoted successful

experience. It is implausible that not working late itself promoted successful experience; rather, it seems plausible that people who strongly disagreed that they worked late on balance possessed some skills or traits unmeasured in the model that both allowed them to not need to work late at the job and that also promoted successful experience.

13. An alternate explanation might have been that people with high education and low job level were predominantly younger, less-experienced employees, while those with low education and high job level were predominantly older, more-experienced ones, and that there were other factors about such employees (separate from their education and job levels) that promoted each group having successful experience. To test for this, job longevity (log scale), and a job longevity quadratic were tested in this model, to see if there might be a nonlinear U-shaped curve whereby employees with little experience and lots of experience, controlling for education/job level, had more successful experience. The results for both variables were not significant (p=.99 and .79 respectively).

19. However, the effect of job level was relatively stronger than education here compared with the first experience model. Furthermore, the effect of job level was higher in the model explaining empowerment than in either the models explaining reducing burdens or getting better value; for reducing burdens, neither education nor job level was a significant predictor of successful experience. These results suggest that part of the impact of job level here was that reform actually did empower people at higher job levels more than those at lower job levels, so part of the result does not reflect job level as a facilitator of successful experience. The insignificant results for successful experience reducing job burdens suggests that cognitive and experiential resources were not necessary for success at reducing burdens, a finding that makes sense.

21. The impacts were somewhat different, but also understandable, for experience regarding empowerment, getting best value, and reducing burdens. For successful experience getting best value, training was significant as a main effect -- greater training was associated with more-good experience with best-value contracting for all respondents. Training provided specific skills for doing best-value contracting that affected a person's ability to be successful. However, for successful experience being empowered, training had a positive impact only for respondents with positive, or at least neutral, expectations about reform. (In other words, there was an interaction between the impact of training and initial expectation in explaining successful experience.) Whether one feels empowered has a greater element of personal reaction to it than the other experience variables. Training made it easier for those with positive expectations to get skills they needed to feel empowered successful experience. For those with negative expectations, training actually contributed to negative experience. Finally, training was not significant at all (either alone or interacting with expectations) in explaining successful experience with burden reduction. Training wasn't really needed for successful experience here, just as cognitive and experiential resources had not been needed either.

23. This difference between the two models may be due to greater impact of the two factors discussed earlier driving the possible asymmetric impact of the self-fulfilling prophecy. At the beginning of the process there was no backlog of experiences where one could look back over time and remember positive ones more strongly than negative ones. Finally, earlier in the change process, supervisor support for reform was lower than later on, so supervisor blocking of the translation of negative expectations to negative experience would have been less strong a factor.

27. This variable should be seen as measuring variation across respondents in the impact of these various mechanisms in producing attitude change. Differences in impact across individuals might be a result of different dispositions. One can certainly imagine, for example, that people might vary, say, in tolerance for dissonant thoughts. The limited research on individual differences in the strength of the drive to reduce cognitive dissonance has produced mixed results. One study did show that people with more dogmatic personalities experienced higher dissonance levels from a given situation. (Sherwood et al l972: 85-88; Wicklund and Brehm l976: Ch. l4) Cialdini et al (1995) developed a specific scale measuring a "preference for consistency" (including questions such as "I want to be described by others as a stable, predictable person" and "I don't like to appear as if I am inconsistent"), rather than trying to test for effects of other personality traits such as authoritarianism. They found that those scoring high on their scale were

more likely to be subject to reduction of cognitive dissonance and foot in the door effects. One can certainly also imagine that, faced with the same situation, different initial skeptics/fencesitters might have different feelings about how voluntary their participation in reform-oriented activities was, which would also affect the strength of the cognitive dissonance mechanism. Additionally, differences in impact could reflect different situations in which respondents found themselves: situations may have been structured to make individuals feel they had greater (or less) choice over the decision to participate in reform-oriented activities, and, of course, the number of times individuals participated in such activities (and thus the strength of the force exerted by the various mechanisms) would vary as well. Finally, differences in the impact of cognitive dissonance or self-attribution across individuals would also reflect the strength (either pro or con) of their initial attitudes towards the change.

31. Furthermore, because of current-attitude bias, people tend to underestimate the extent to which they in reality liked the changed ways of behaving more now than in the past (the bias produces a situation where some overestimate the extent to which they liked these ways of behaving as well in the past as they do now). Thus, reported responses understated the extent to which respondents actually did like the new ways more now than before, and thus the coefficient for this variable.

These results are conservative for a final reason. Acting early in the change process to increase attitudinal support for change, these mechanisms contributed to creating expectations for positive experiences that, through the self-fulfilling prophecy, increased positive experience (measured by the other variables in the model such as empowerment and getting best value). Thus, the "as time goes by" support partly explains values for the other variables in the model through an internal causal structure among predictor variables, reducing the reported effect of this variable in the model.

32. This figure is too high because, by eliminating the control variables, one is measuring both the indirect (via experiences) and direct impacts of these organizational influences. However, this figure does represent the total feedback influences coming from an individual's organization, both indirect and direct. And it should be remembered that the effects of experiences on attitudes with which we are comparing the effect of organizational influences is too high as well, because it includes the total impact of experiences on attitudes, both the part of a respondent's experiences caused by its "inherent" features and the part itself caused by feedback influences.

34. The coefficient (slope) for mean workgroup attitude at each value mean workgroup attitude took in the sample was calculated using differential calculus based on the polynomial function involving mean workgroup attitude. The average coefficient for mean workgroup attitude was then calculated by taking the weighted number of observations for each value of mean workgroup attitude, assigning the appropriate coefficient to each value, summing the coefficients, and dividing this sum by the number of observations in the sample. Since mean workgroup attitude also interacted with mean workgroup attitude standard deviation, the coefficient for each value of mean workgroup attitude took into account the actual value of mean workgroup attitude standard deviation associated with that value. One may then use that average coefficient to calculate the impact on reform attitude of, say, a ten-unit change in mean workgroup attitude.

36. 13% of the sample had values of mean workgroup attitude standard deviation that were less than or equal to 10, the mean standard deviation was around 20, and 11% of the sample had values greater than or equal to 20. The lowest value for mean workgroup attitude standard deviation in the sample was 5.8 and the highest 37.5. However, the workgroups with these extreme values were very small.

37. So, if one compares the impact of mean workgroup attitude on individual support for reform where mean workgroup attitude standard deviation is assumed hypothetically to equal 20 at all values of mean workgroup attitude (column two in the table) with the impact given <u>actual</u> sample values for mean workgroup attitude standard deviation at different levels of mean workgroup attitude (column four), one sees that in the latter case the asymmetric anti-reform impact of mean workgroup attitude was less dramatic than in the former.

39. To the extent a respondent might have chosen a most-respected coworker because of the coworker's agreement with the respondent's pre-existing views on reform and/or (in the context of the variable keeping up with leader statements, to be discussed below in the section on leader influence on attitudes) begun or stopped listening to statements of top procurement leadership because the respondent found the statements of pro-reform top leaders more or less attractive, this would produce endogeneity -- the direction of the causal arrow between most-respected coworker attitude/keeping up with leader statement and support for reform would be reversed, meaning that reported coefficients would be exaggerated and coefficients for predictor variables correlated with most-respected coworker attitude/keeping up with leader statement would be too low. It should be noted that the statement in keeping up with leader statement was presented as a general disposition, not tied to specific top leaders ("I try to keep up…"). I would also add that I am skeptical that many would change the person whose opinion they respected the most because of the person's views on procurement reform.

It is also possible that another source of endogeneity might exist -- the association between most-respected coworker attitude and support for reform could be partly due to the respondent influencing the attitude of his or her most-respected coworker, rather than the other way around, as is being assumed here. (I am grateful to my colleague David Lazer for pointing out this potential problem to me.) I would suspect that, psychologically, when somebody chooses another as "most-respected," he or she imagines somebody who influences them rather than somebody they influence.

Results reported earlier suggest endogeneity was not a severe problem. As regards most-respected coworker attitude, there was no relationship between having a most-respected coworker who was an earlier supporter of reform and the respondent's own earlier support for reform, except for members of the change vanguard, a small group for whom such a choice pattern probably did occur; if people had been choosing their most respected coworker based on congruence with the respondent's own attitude towards reform, a positive relationship should have existed. (It is possible, of course, that such selection might have started after reform began, which the earlier results would not have captured.) As regards keeping up with leader statement, prior to the beginning of reform, people (in civilian agencies) who kept up more with top leaders agencies were less likely to join the change vanguard, which is inconsistent with the view people choose to listen to leaders in order to hear pro-reform ideas. Also, people with high scores on keeping up with leader statement reacted differently before l993, when the attitudes of the top leadership changed, and after l993, a result inconsistent with a view that answers to this question merely reflected a tendency after l993 of pro-reform respondents to begin for the first time to listen to pro-reform top leaders, as opposed to an underlying propensity to listen that was constant before and after l993. It is possible, however, that there was some endogeneity, which would have created a general overlay on the results from Chapter Three just discussed.

42. Note again the contrary finding in the Damenpour meta-analysis reported discussed in a footnote to Chapter Three. It may be noted Damenpour's simple variable captures nothing about the extent to which leaders seek to influence opinions or about leader persistence. The studies also seem to involve mostly top leaders, so they wouldn't capture either behavioral facilitation or direct attitudinal influences of lower-level leaders.

52. T o make these respondents as similar as possible to ones who left, I limited the comparison to respondents who had started working in government procurement before l993 and had been at their current office before that year as well.

54. There were a significant number of respondents who left the question blank regarding what year they had become a supervisor, so the numbers here are smaller than the total number of supervisors in the sample. Team leaders were also excluded from these data.

Chapter Nine

2

8. The emphasis on situational rather than attitudinal factors in explaining behavior corresponds to the general trend in social psychology towards emphasizing situational over "dispositional" factors in explaining behavior. See, e.g. , Ross and Nisbett (1991). It should also be noted that economists have traditionally shown little interest in attitudes for a somewhat different reason. Ever since a classic article by Samuelson (l938), economists have tended to look at choices (i.e. behavior) and ignore preferences (i.e. attitudes), referring to choices as "revealed preferences." An important part of the reason for this has been methodological – the view that inner mental states cannot be observed, but behaviors can – although the replacement of an analysis of preference or utility by choice has also had theoretical justifications (that need not concern us here) as well. (Sen 1982: 54-73) One may note, however, that, although economists share with attitude-behavior skeptics a distaste for examining attitudes, the view that choices are "revealed preferences" actually suggests a close relationship between attitude and behavior, i.e. preference and choice, rather than the non-relationship the skeptics espoused, since the idea is that (observable) behaviors "reveal" unobservable preferences.

11. Furthermore, attitudes arising out of direct experience are more readily accessible to a person at the time of action compared with those not so based. The strengthening effect of direct experience on the attitude-behavior link suggests that attitudes about certain kinds of situations or objects are likely to be more closely linked to behavior than others.

13. Brief discussions appear in Fishbein and Ajzen l975: 344, Schuman and Johnson l976: 198-90 and Eagly l992: 695. Eagly notes that interaction effects "have for the most part not been incorporated into explicit theories of attitude-behavior relations." Songer-Nocks (1976: 67) discusses interaction effects between attitude and prior experience with a behavior in explaining the attitude-behavior relationship. A very small number of studies have looked for, or been reanalyzed to check for, interaction effects involving an individual's attitudes and the attitudes of a reference group salient to the individual. (Liska l974)

This lack of discussion is particularly surprising because many discussions of ways the attitude-behavior connection gets blocked describe interaction effects, albeit in words. The more recent literature asks "under what conditions" attitudes get translated into behavior, and examining interaction effects allows answering that question, in a way analysis only using main effects cannot.

24. There has been some limited attention in the labor economics literature to the research on intrinsic motivation (e.g. Kreps l997 and Prendergast l999: 19). Kreps (ibid.: 360), observing that he co-teaches a course in human resources management at a business school with a sociologist and a social psychologist, notes that doing so requires that he "assert things that, as an economist, I do not really understand," such as that "providing extrinsic incentives for workers can be counterproductive." In a brief discussion, Prendergast states that "this idea holds some intuitive appeal," but notes that experiments showing that provision of extrinsic incentives reduces voluntary continuation of the activity after the incentives are withdrawn may simply reflect the fact "that if those who operate on piece rates perform better during the experiment period, they are simply more tired of carrying out that activity than those who have operated at a ore leisurely pace without pay for performance. Thus they may be less likely to continue the activity for reasons other than intrinsic motivation." However, studies have been designed to obviate any satiation effect and still show the same results.

Note that many of the jobs, such as windshield installers or sales clerks, examined in economic studies of productivity effects of economic incentives, provide little intrinsic motivation to those performing them. The empirical literature in economics also concludes that a substantial portion of the productivity effects of extrinsic rewards comes via a selection effect (i.e. people believing ex ante they will be below-average performers don't seek jobs with this reward structure in the first place, or leave when they are introduced) rather than a motivation effect.

Alternatively – and more intuitively for economists – one may argue that extrinsic rewards changes the total utility from the behavior: previously the utility derived from chess was enjoyment at the game plus opportunity to choose and demonstrate competence, and extrinsic rewards remove the perception of choice, lowering total utility. (This is the approach taken in Gneezy and Rustichini 2000.)

In principle, in an empirical test for the undermining effect hypothesis where one's underlying attitude (in this case one's attitude towards procurement reform) was in the model, any attitude change due to the "contamination" of extrinsic rewards would be expected already to have been measured in the respondent's report of his or her attitude, which were therefore take such contamination into account. In practice, though, the data are surely too blunt to detect whether contamination is occurring. The feeling thermometer question simply measures attitude towards "procurement reform," not "procurement reform because it's the right thing to do" versus "procurement reform to please the boss." Possibly, the rating a respondent gives for "procurement reform" already reflected, in the case of an attitudinally sympathetic respondent with a supportive boss, any attitudinal "discount" (so to speak) arising from the respondent being in a situation where attitudinal support for reform was taking place in the context of a supportive boss, in which case any decline in the expected behavior changed of the intrinsically motivated has already been reflected in a decline the self-report of the respondent's attitude that this phenomenon had produced. Ratings respondents gave may also have been more abstracted from the respondent's particular environment and thus did not reflect such a "discount." Most likely (especially since respondents surely didn't consciously reflect on this distinction in answering the question), there is some of both.

26. The picture becomes more ambiguous because rewards may also provide information that a person is competent at performing a behavior; since the theory postulates that the opportunity to demonstrate competence is a source of intrinsic motivation, the more a reward provides such information (say, by being contingent on successful performance), the more it might promote intrinsic motivation. So, for example, what Deci and Ryan (l999: 628) call "engagement-contingent rewards," rewarding people for engaging in an activity without requiring successful performance, provide no information about competence and thus nothing to offset the depressing impact on intrinsic motivation of their being controlling. By contrast, "performance-contingent" rewards, rewarding people for meeting or exceeding a certain standard, can also convey substantial positive competence information in cases where the person does well enough to get a level of reward that signifies excellent performance. In those cases, there is a significant tendency for performance-contingent rewards to affirm competence and, accordingly, to offset some of the negative effects of control. (Ibid.: 629)

The net effect of performance-contingent rewards on intrinsic motivation would then be theoretically indeterminate. Empirical studies (summarized in Ibid.: 643-45) generally show that, on balance, they reduce intrinsic motivation, i.e. that the control effect outweighs the competence-provision effect.

31. The attitude quadratic term also allowed controlling for an alternative explanation for significant interactions between reform attitude and a third variable in explaining behavior change. Since an interaction is a multiplicative term, like a quadratic term, what appears as a significant interaction between attitude and a variable correlated with attitude might in effect be a quadratic term for attitude itself, simply reflecting that attitude's relationship to behavior change is curvilinear. (I would like to thank my colleague Dani Rodrik for discussing this issue with me.)

32. These findings control for cognitive distortion that might lead reform supporters to overestimate the degree to which they had changed their behavior and reform opponents to underestimate it. They do not take into account, however, the likelihood that some of the relationship between attitude and behavior involves behavior causing attitude, rather than the other way around, which would make the reported link exaggerated.

33. 6% of the sample had a value for reform attitude less than or equal to 40; 22% of the sample had one greater than or equal to 90. This reflects the predominance of reform supporters.

35. An upper-bound is provided by the first-order correlation between attitude and behavior change (the statistic in most earlier attitude-behavior studies). This is .4, at the top end of the range of those studies, corresponding, in a simple linear regression, to an R-squared of .l6. However, this exaggerates the influence of attitude, because it attributes to attitude all covariance between it and other variables influencing behavior that are correlated with attitude. By contrast, a very conservative estimate is provided by looking at unique variance attributed to attitude. A model run using all variables used in the two models as main effects, but excluding attitude, yields an adjusted R-squared of .l7, while running the same model with attitude produces an R-squared of .27 -- so adding attitude at the end produces a 10 percentage point improvement in variance explained. This is conservative, because it attributes all covariance to variables other than attitude. Also, because it is impossible to parse variance explained by one of two interacting variables, this includes main effects of attitude only, without interactions explaining additional variance.

There are a priori reasons to suspect that endogeneity might be less important, or unimportant, in the case of the relationship of attitude to behavior than in the case of the relationship between self-reported personal experiences regarding procurement reform and reform attitudes. One reason is that a perception that one has personally experienced a great deal of behavior change does not imply happiness or unhappiness with the change in the same way that a perception that reform has made one's job easier (almost certainly) is a positive statement about one's perceived personal experience. An attitudinal critic of reform could easily perceive big changes in the way he or she did his job that were seen as negative. It also seems intuitively plausible to believe that a person who likes reform develops an exaggerated view of how positive his or her personal experiences of reform have been, it seems less intuitively plausible to assume that similar distortion processes would affect the quantity of change the respondent experienced. Finally, to reduce any impact of thinking about one's attitude towards reform influencing a respondent's perception of how much one's behavior had changed, the reform behavior question was posed in the survey before any of the attitudinal questions.

42. Another possible explanation involves respondents' own perceptions of how much their behavior had changed, as opposed to an independent observation. The greater the mean workgroup support for reform, the greater the extent of actual behavior change in the workgroup. Perhaps, the more behavior change people saw around them, the more they tended to undervalue the extent of changes they had undertaken themselves and/or underestimate the extent to which their current behavior departed from previous behavior (since their new behaviors didn't stick out in contrast to others in their environment). To the extent such a perception explanation is correct, this implies more generally that self-reports of behavior change were biased downward and actual behavior change was greater than reported.

50. These results seem inconsistent with dragalong, which can't explain differences in the behavior of respondents equally situated in terms of personal attitudes and the mean workgroup attitude but where there are different levels of workgroup consensus. If anything, one would predict that more-united workgroups, at a given mean level of attitudinal reform support, would be more likely to drag reluctant individuals along with behavior change than less-united workgroups. However, the results were the opposite.

In addition, as only the undermining effects theory predicts would be possible, there was a crossover effect where high workgroup support for reform produced a positive effect on behavior change among reform critics (influenced only by extrinsic rewards) but a negative one for supporters (where undermining occurred). However, this finding is more subject to question than the comparable result for supervisor attitude presented above, because of the general bias in the model for negative coefficients for mean workgroup attitude, which could be producing these coefficients among supporters.

This would most straightforwardly be the case if these other supervisor influences produced behavior effects similar for supporters and critics. One could imagine, though, that various of these other supervisor effects did not act equally at all levels of reform support. In principle, for example, behavioral facilitation might promote behavior change more among supporters than critics: supporters would take advantage of opportunities behavioral facilitation provided, while critics foot-drag and make the minimum change they can get away with. If this were the case, the undermining effect in these results would be understated, because there would be weak or negative supervisor influence on behavior change for pro-reform respondents even after the disproportionate lift to behavior change coming from behavioral facilitation. However, one could also imagine the opposite as well. For reform-oriented activities directed by supervisors, supporters might change anyway, without help, while critics would find themselves forced to participate in activities bosses were directing. If this were the case, the undermining effect would be overstated, because part of the strong effects observed for reform critics would be due to disproportionate influence of supervisor direction on them.

It should be noted, then, that the pattern of results here, including the slope of influence change from critics to strong supporters, is very similar to that seen in the interactions with outside pressure and centralization, neither of which are affected by behavioral facilitation or supervisor direction. This is consistent with the view that these influences act similarly at different levels of reform support.

Chapter Ten

1. For procurement people buying weapons systems, better cooperation with industry might be seen as reducing job burdens to the extent that it reduced requirements for industry to submit, and the government to analyze, industry cost data for the purpose of pricing what the government was buying.

8. These were respondents stating earlier in the interview that they regarded the traditional system as "broken" but who, asked why, had not mentioned any criticism involving failure to produce better value. This is a somewhat broader group than Those who were initially burden-reducers, because it includes many criticizing the failure of the traditional system to provide job autonomy, not just those seeking burden reduction.

9. This may exaggerate somewhat the percentage showing deepening. Some might have from the beginning seen the reform effort as involving the better value agenda even though they themselves didn't embrace it. And this question only required respondents to give their perceptions of the general goals of reform, not to endorse them, so a respondent might have perceived better value to be a reform goal while personally being uninterested in it.

28. The greatest difficulty was to distinguish between successful experience and "as time goes by" support. For the former, I looked for mention of success with experiences; if the response didn't mention success but still reported attitude change, this was coded as the latter. The In-Person Interviews were not coded in the standard manner using independent coding by two coders (for a discussion of why, see "Sampling Issues and Survey Administration"; this is one question where there might well have been poor inter-rater coding reliability. This is another reason to regard these answers as suggestive.

The numbers presented here may exaggerate somewhat the percentage showing deepening. Some might have from the beginning seen the reform effort as involving the better value agenda even though they themselves didn't embrace it. And this question only required respondents to give their perceptions of the general goals of reform, not to endorse them, so a respondent might have perceived better value to be a reform goal while personally being uninterested in it.

29. Respondents in non-military organizations were eliminated from the regressions for the milspec variable, and respondents in organizations buying weapons systems were eliminated for the performance-based service contracting variable.

31. I also ran the model with the question, discussed earlier in a different context, about procurement reform making it easier to select best-value contractors. Results with this other variable in the model (not shown here) were virtually identical; the few differences will be noted. However, this variable was somewhat problematic as a measurement for the impact of positive experiences involving the best value agenda for a different reason, since it addressed only one feature of that agenda, namely selection of contractors providing the best value (such as through the use of past performance in vendor selection), not

other elements, such as experience with milspec reform or performance-based service contracting specifically, and innovativeness in general, making results with this specification conservative as well.

There was also evidence that the more appropriate (i.e. least overinclusive) the variable used to measure successful personal experience with the better value agenda, the larger the importance. For past performance, successful experience was significant (p=.0001) when specified using getting best value rather than successful experience, for which it was insignificant. Given the importance of past performance in new methods for selecting good contractors, this would be the kind of successful experience relevant to the better value agenda most relevant here. Similarly, successful experience had both a high significance level and effect size predicting support for milspec reform. I would conjecture this was because for many Defense Department respondents, milspec reform was a central feature of their encounter with procurement reform, and responses to successful experience in significant measure were equivalent to experience with milspec reform.

Getting best value was significant at the .l level in predicting support for cooperation with industry, but the sign was in the opposite direction – positive personal experience getting contractors offering the best value was negatively associated with support for partnership with industry. This may be due, among this subset of respondents at least, to anti-contractor attitudes that might be causing both a view reform allowed "toughening up" on poorly performing firms and a hesitancy for too close a relationship with contractors.

41. Other, albeit statistically less strong, evidence for the same argument comes from a regression equation using the different types of successful experience with reform as predictor variables, with categorization as the dependent variable. Comparing this regression with the same model with reform support as the dependent variable shows different results. In the model in Table Two, both making one's job easier and empowerment had been significant, and indeed the job easier question had a weaker effect than empowerment. Xx However, where categorization was the dependent variable, the effect size for making one's job easier was the largest in the model, and that for empowerment was trivial. (It should be noted, however, that the p-value for the job easier question was .l8; the result for the success story question was statistically significant, but its effect size was smaller than for the job easier question. This again suggests that many factors might produce reform support, but burden reduction was central to producing categorization.

44. It will be remembered local leader attitude took only two values, "l" for a member of the change vanguard who had been at the job since the beginning of the reform effort, and "2" for everybody else. No local office heads in the sample were actively hostile towards reform; this means this variable (since all such people, as reported earlier, supported the better-value agenda) reflected variance in degree, and length of time, of support for the better-value agenda.

The same specification for individual supervisor attitude was used as in the earlier model predicting deepening attitude, and the mean workgroup value was computed in the same way as for the mean value for the dependent variable.

Chapter Eleven

10. Given that persistence is seen here as a system-level variable (that thus doesn't vary across our sample), and we have only one case, how can we know persistence made a difference? The argument is partly a logical one. If it is true some kinds of positive feedback increase with mere passage of time, and if in other cases feedback mechanisms only turn positive in net effects relatively later in a change process, then the longer an effort lasts, the greater the effects of positive feedback are likely to be, unless any decline in the impact of positive feedback mechanisms depending on mere initiation outweighs increased impact of other forms of positive feedback later on. Also, one may make casual empirical comparisons (such as the ones above) to situations where persistence was and was not present.

Appendix A

1. It is possible that both education and job level are ordinal rather than interval variables – i.e. for example, that the impact of a move from a high-school to a college education produces larger results than the move from a college degree to a graduate degree. One could test this (at least in ordinary least-squares regression) through a quadratic term testing for curvilinearity or through testing for different coefficients for dummy variables at different values of these variables. These distinctions are not important to this analysis, and we will assume throughout these are interval variables.

For non-supervisory contracting officers and supervisors, respondents were asked what year they had attained their current position; for this model (and elsewhere in this chapter where this variable is used), those who had been promoted into their current jobs in l993 and after were recoded to one level below their position at the time of the survey. One difficulty in this recoding is that "team leader," one of the job level categories, was largely a new position created after l993 in response to the reinventing government mandate (applying to agencies in general, not just procurement) to reduce supervisor-to-non-supervisor ratios. (Team leaders don't perform performance evaluations and hence are not counted as supervisors.) In the Frontline Survey, 82 of 89 respondents classifying themselves as "team leaders" reported they had become team leaders after l993. The problem is that some team leaders were promoted from non-supervisory contracting officers, while others were demoted from supervisory positions, so the data I collected did not allow me to assign these respondents a "former" position. I kept these respondents as an intermediary position between non-supervisors and supervisors, as they were in the survey.

2. These questions were selected from other survey research measuring risk-tolerance and were chosen so as not to involve job-related situations in order to tap an underlying personality trait. Initially, I assumed that venturesomeness and risk-tolerance captured the same phenomenon and intended to include in the same scale another question, "I try to avoid possible cancer-causing substances in the food I eat." It turned out, however, that responses to this question were very weakly correlated with responses on the other three (.08). The cancer-causing substances question thus seemed to be measuring the willingness to take a chance of loss in case of failure in order to gain the benefits of success, rather than venturesomeness. One could certainly argue that the question about climbing mountains includes some element of risk-tolerance as well as venturesomeness, which would put at least some measure of risk-tolerance into this scale.

Because of the effort with the scale to find some underlying personality trait across various domains of venturesomeness and risk-tolerance, not confined to one kind of cautiousness, respondents needed to answer at least three of the four questions to be given a value for this scale. Others were coded as missing values. It may be noted that a number of experiments have failed to find strong, or even significant, correlations between a subject's risktaking propensity in different kinds of laboratory situations, a conclusion in line with the general view in social psychology that situation-specific factors influence behavior more than underlying personality traits. (See the early experiment by Slovic l962 and the literature summary in Slovic l972.) However, in one detailed examination of risktaking among managers, participants provided responses to various standardized situations involving on-the-job problems, investment decisions, and gambling, as well as to behavioral questions regarding insurance holdings, debt, asset distribution, gambling activities, and hazardous recreational activities. (MacCrimmon and Wehrung 1986) This study found that, across the various situations, there were more than five times as many consistent risk-averters as would have been predicted by chance and three times as many consistent risktakers. (Ibid.: 204) To the extent that caution is not a stable personality trait across situations, results here would be expected to be conservative, since the questions in the scale are far removed from on-the-job experiences, so they may underestimate connections between job-related cautiousness and earlier support for reform.

A number of the scales to be presented in this and other chapters have a Cronbach's alpha below the normally desired level of .7. One reason is that the length of the survey certainly increased the proportion of random answers, producing more noise, which reduces the internal structure of a scale and lowers the alpha coefficient. Second, these scales virtually all involve questions measuring psychological dispositions. My guess is that, in a survey presented as being about procurement, a significant number of respondents were puzzled by or resentful of questions having nothing to do with procurement, such as whether one liked to plan trips carefully or respect the authority of management, producing greater randomness in answers and hence answered these questions randomly because they resented being asked to answer them at all. One piece of evidence for this explanation is that generally a much larger percentage of respondents (often around 15%) left these questions blank, or responded "don't know," than for the more procurement-related questions; one way to deal with a question one doesn't want to answer is by leaving it blank, while is to answer randomly. By contrast, the Cronbach's alpha for the one scale using questions related to the respondent's job in procurement, on job burdens, did have an acceptable alpha coefficient of .70. The alpha coefficients to be presented certainly suggest some significant structure to the scales. The presence of noise suggested by low coefficients makes it harder for the relationship between the variable in question and the dependent variable to attain significance, and hence makes these results a conservative test of significance.

4. The dependent variables both measure a respondent's current attitude, not the attitude before the beginning of reform. It may be presumed some shift in opinion towards greater desire for job autonomy as a result of reform occurred. However, although the shift affects the mean value of this variable, there is reason to believe, based on analysis in public opinion research on the impact of forces that tend to shift opinions over time on different subgroups in a population, that this shift will distort neither coefficients nor statistical significance levels in this model. This would also be expected to be the case for other current-report variables appearing in this and other models in this chapter. Page and Shapiro (l992: Ch. 7) study whether exogenous factors causing shifts in public opinion on a question differentially affect different population subgroups with different initial attitudes on the question – e.g. if an event, such as the September ll, 200l, terrorist attack, or a social trend, such as decline in racism among whites, differentially affects those who previously had been low and high in patriotism, or low and high in racism. The authors define a differential trend as one where the post-event gap between the opinion of subgroups was l0 percentage points or more greater than the pre-event gap. Using this measure and examining 3000 studies, they found that in only 5.7% of the cases did opinion, measured as percentage point gap between the attitudes of subgroups, diverge over time. And virtually all the subgroups they measured were demographic ones (religion, age, race, education) where there is actually more reason to believe exogenous forces would affect groups differently; for the one opinion-related subgroup, dividing respondents by partisanship, only 4.3% of cases showed divergence over time. (Ibid.: 290-91) "Most of the time," Page and Shapiro conclude (Ibid.: 290), "the policy preferences of different subgroups changed (or stayed the same) in similar ways. The trend lines were mostly parallel." Indeed, the authors call their chapter "parallel politics." (On a related issue – the fact that British election analysis typically refers to, and observes, a uniform percentage "swing," over time and in response to exogenous events, across constituencies with widely different initial voting patterns -- see Butler and Stokes l97l: 86-94.) Translated into this context, if we divide respondents up (to simplify) into two groups, "initial job-reform skeptics" and "initial job-reform backers," we can expect that, if the percentage of respondents scoring high on disliking rules would have been, say, 30% if we had measured the variable contemporaneously, for job-reform skeptics and 60% for job-reform backers, the prediction would be that, as changes due to procurement reform would increase the percentage scoring high in both categories by about an equal percentage. Translating the equal percentage into mean values, this would mean there was an equivalent increase in the mean value for disliking rules at both levels of initial job-reform support, meaning that the observed distribution of values for disliking rules based on the current-attitude report would simply shift the intercept of the variable upwards, without changing its slope, meaning no effect in the model on coefficients or statistical significance levels of predictor variables. I am grateful to my colleague Suzanne Cooper for her advice on thinking through issues of possible biases created by use of current-attitude questions in models involving an earlier time period.

For the retrospective recall predictor variables in this model, the Frontline Survey comparies current with five-year earlier perceptions. In all cases, they show respondents regarded pre-reform problems as having decreased – they thought, for example, the system provided more timely service currently than previously and that contractors currently performed better than previously. This suggests retrospective reports would tend to underestimate problems in the traditional system, creating a current-attitude bias, which tends to bias coefficients for these variables upwards, for reasons discussed in "The Impact of Sample Attrition and Retrospective Recall Bias."

However, it is also possible a person's answer will be biased in a "good old days"/"bad old days" direction based on their current views of reform, which is the opposite of a retrospective recall bias. In this case, anti-reform people will remember the previous period as "better" (e.g. less stressful) than they actually would have experienced it at the time, so reported results underestimate previous stress, while pro-reform people will remember the previous period as "worse" (e.g. more stressful) than they actually would have experienced it at the time, so reported results overestimate previous stress. That would create the opposite bias.

Even if variables measuring attitudes towards job autonomy at the time of the survey probably do not affect coefficients in the models, it is possible there may have been a shift over time in reasons why people sought or shunned job autonomy – a change in the relationship between predictor variables and the dependent variable. In particular, efforts during the years of the reform effort to encourage people to be innovative and to find new ways to do business better may have produced an increasing tendency to support autonomy in order to achieve better organizational results. To the extent this occurred, coefficients for better-results variables exaggerate the extent to which they were driving attitudes towards job autonomy before reform began.

Finally, in subsequent models, I will create a separate variable for respondents with missing values. However, doing so in EQS created significant problems for the goodness of fit statistic of the model, because of interconnections among these missing value variables. For EQS, therefore, I imputed a value for missing values (both for predictor variables and the dependent variable) based on the respondent's value for reform attiude, which, as I explain below, non-randomly varied among respondents and non-respondents to questions.

5. Strictly speaking, this doesn't measure how much one cares about making good decisions, only an empirical view about the connection between one's performance and agency success. However, given the statement's wording, it's hard to imagine very many agreeing wouldn't care about how well they did their jobs. (To do so, one would need to say, "How I do my job has a big impact on how well our agency performs, but I still don't care how well I do.")

7.One must control for job level, since an apparent positive effect of the desire to succeed on the desire for job autonomy might simply reflect the fact that people driven to succeed are more likely to be promoted, and people at higher job levels are more likely to seek autonomy.

Both job autonomy/disliking rules and being driven to succeed are measures of current attitude. After the spread of reform, which promoted job autonomy, a person driven to succeed at his or her job might, as of the time of the survey, have expressed more support for job autonomy because doing so might gain favor with the boss, meaning that reported values for disliking rules/job autonomy for people who strongly agreed they were driven to succeed might have been too high. If this was occurring, this would mean that the reported slope of the relationship between being driven to succeed and disliking rules/job autonomy was too steep, because the change in y (disliking rules/job autonomy) produced by a change in x from a lower to higher desire to succeed was greater than it actually would have been at the time.

8. Also, without a measure of tenure, a significant association between age and support for autonomy (the "postmaterialism" hypothesis) might reflect omitted variable bias -- one could argue the observed relationship was occurring because young people had been exposed to the traditional system for less time, not because they were more affluent. Having job tenure in the model also deals with potential omitted variable bias involving education (without controlling for tenure, one could argue the apparent link between education and support for autonomy was actually occurring because people who had started working a long time ago in procurement, who opposed autonomy because of commitment to the traditional system, also had lower education, because people used to start with lower education levels).

10. A question measuring association of dissatisfaction with timeliness –an empirical statement – with earlier support for reform was thus used to measure whether support for TQM – a statement about one's

values – drove earlier support. This is legitimate because if a person didn't think timeliness was important in the first place, it's hard to see how failure to provide timely service could drive earlier support for reform.

12. One question mark should be noted about interpreting this variable. One of the substantive tenets of reform was that procurement people should become more willing to take risks in decisions they made. It might therefore be the case that the strong effect of this variable reflected the extent to which respondents were oriented towards taking substantive risks when they made decisions. One piece of evidence against this alternative interpretation of the data is that, as noted earlier,. the venturesomeness/risk-tolerance scale did not correlate strongly with current support for reform, which one would expect to have been the case were the scale influencing reform attitudes via a substantive emphasis on risktaking (at a much later time in the process, where earlier support was no longer an issue).

18. This is a current-attitude variable. It is possible that, compared with the period when reform was beginning, initial reform supporters had become more likely to keep up with leader statements because top leaders were now saying things they found more attractive. If that is the case, some respondents reporting as of the time of the survey that they kept up with top leader statements a great deal would actually as of the beginning of reform have been keeping up with top leader statements less, which would mean that the coefficient based on current-attitude reports was too low. (To illustrate with hypothetical numbers, if the mean score on this variable for initial supporters was 2.0 and the actual score based on the extent to which they had kept up at the time was 2.5, it would have taken a smaller change in x to produce a given change in y, from greater to less initial support for reform.) However, if people changed the extent to which they kept up, or didn't, with top leader statements based on support or opposition to reform, this would imply some endogeneity in this variable, which would make reported coefficients too high. I will return later to a discussion of possible endogeneity for this variable.

The question asking respondents whether they read Federal Contracts Report was a current-attitude question. To the extent that people started reading FCR after l993 because they wanted to read coverage of procurement reform, this would bias results downward.

16. It is of course possible that rebelliousness might reflect a motivated desire to be different or to try different things, rather than a resentment of authority that simply assails the person, which would imply that rebelliousness in fact might be associated with a desire to innovate on the job. However, a variable measuring the desire for original thinking is also in the model, so this controls for that possible impact of rebelliousness.

18. This was felt to be an unthreatening way to ask a respondent about the extent of unease about trying new ways of doing the job; it was also less subject to worries about direction of causation from a question inquiring directly about the extent to which change was initially seen as stressful. Skepticism about the specific changes reform involved, rather than stress about job changes in general, might have caused stress in this particular situation. However, the coefficient for this variable might still be exaggerated by reverse causation: the extent to which one felt less sure one was doing the change the right way might be influenced by one's initial attitude towards the change.

19. This coefficient is likely to be too low because of sample attrition.

19. I used this variable to predict resistance to change associated with organizational exposure in particular, rather than age, which might measure resistance to change in general. Results putting age in the models instead of this variable turned out to be the same.

20. Originally, I developed five categories, ranging from enthusiastic support to clear opposition to reform. These categories would almost certainly have had to be seen as an ordinal scale, which is not suitable for ordinary least-squares regression analysis. I experimented in running the regression models with both this five-point scale and with various other recodings to see if some might work better. The one that worked the

best separated the highest category in the original six-point scale – local office heads who were both the very most enthusiastic and also had served in their jobs a long time -- from everybody else. 32% of the weighted sample had leaders in the "most-enthusiastic" category. One might imagine that if there had been strong opponents as well as a range of supporters of change, any coefficients would have been larger. Furthermore, of the l9 organizations, for the one whose office head was closest to a critic of reform (the only one to classify herself as having "mixed feelings about acquisition reform"), we do not, as noted earlier, have any meaningful number of responses to the Frontline Survey. Nonetheless, I would certainly perceive variance among these local office in the degree of ardor they showed for reform (although, to be sure, ranging from moderate support to enormous enthusiasm); furthermore, some of these leaders had arrived relatively recently, while others had been in their organizations since (or before) l993.

21. This question referred to the respondent's supervisor at the time of the survey. However, the survey also contained a question about the attitude of one's <u>previous</u> supervisor if the respondent had gotten a new supervisor over the previous two years. For respondents reporting a new supervisor, the response for the attitude of the previous supervisor was used.

It is possible that this self-report was distorted by the respondent's own opinions about reform, i.e. that a pro-reform respondent misperceived the attitude of his or her immediate supervisor as being more pro-reform than it actually is, while an anti-reform respondent has the opposite misperception. Since the dependent variables in these models correlate positively with attitudes towards reform, such distortion would create the risk of endogeneity – a situation where the causal arrow in fact proceeded from dependent to the predictor variable, rather than the other way around, as we are trying to test here. Where there is endogeneity, the coefficient for the predictor variable is too high (it reflects y causing x as well as x causing y), and coefficients for other predictor variables correlated with the endogeneous variable are also distorted. Econometricians generally try to deal with this potential problem using two-stage least squares regression, where one seeks to develop, if possible, an instrumental variable that allows for a test of the extent to which x indeed <u>caused</u> y. In this case, with no appropriate instrumental variable available, I tried to deal with this problem by developing a variable measuring distortion, that works to control for such misperception, that will therefore be included in these models. I will discuss this issue further, including presenting how the variable was developed, in Chapter Five, because the problem is larger in looking at the relationship between various perceptions of the effects of reform and one's current attitude towards reform. The risk of endogeneity will arise in connection with a number of the predictor variables used in various of the models to be developed in this and subsequent chapters. Where this is an issue, I will note and discuss it in footnotes, including discussing the extent to which one has reason to suspect the problem is likely to be large or small. I was not able in any cases to develop instrumental variables that would have permitted two-stage least squares analysis.

6l% of respondents reported getting a new supervisor over the previous two years. Given this high turnover, a relatively significant number of respondents whose supervisor in l993 was not the same as their supervisor two years prior to the survey (1996 or l997). If supervisors leaving between l993 and l996-97 were more negative to reform than the supervisors for whom we have data here, then the location of the regression line shifts to the right (the real distribution of supervisor opinions is less pro-reform than the observed one), but the slope of the line doesn't change.

22. This is also subject to the same perceptual distortion as supervisor attitude, though for the purpose of analysis in this chapter, I don't believe this constitutes as serious a problem, given the way this variable is coded -- there is no reason for potential distortion to extend to an observation that the opinion leader was an <u>earlier support</u>er of reform, which is the response focused on here. However, the distortion variable in this model corrects for this potential problem as well as the one with supervisor attitude And, as we shall see in the data analysis in this and the subsequent chapters, a respondent's answer to most-respected coworker attitude and supervisor attitude have very different effects on predicting different variables involving different aspects of support for procurement reform, suggesting these self-reports tap something real and are not just a common artifactual correlation of answers with support for reform. Given the possibility that somebody who changed offices since l993 had also changed most-respected coworker, as with local leader, those who had done so were coded as missing values.

23. It will be remembered that we knew the buying office for each respondent. We could therefore compute the mean value for this variable at a workgroup level. For 69% of respondents, information was also available about the division – the level below the buying office level, but generally one or two levels above the primary workgroup level headed by a first-line supervisor. For respondents for whom division-level data was available, the mean value at the division level was calculated; for others, the mean value at the buying office level was calculated. It should be noted that in neither case was data available at the primary workgroup level, where influence might be most likely to be exercised. This means that these data are noisy and that p-values and coefficients will be conservative, underestimating impacts.

In calculating the value for this variable, the following respondents were removed from the denominator and the numerator: (l) the respondent himself or herself, so that results would reflect only influence from others in the workgroup (this meant that the value might be different for different respondents in a workgroup), (2) those who had worked in government procurement in l993 but not at the same office where they were working at the time of the survey (the Frontline Survey included a question asking respondents, "What year did you start working at the office you're now working at?"), (3) respondents with missing values for initial attitude. Since some respondents might have shifted workgroups since the start of reform to join one more sympathetic with their views, if the respondent himself or herself had not been in the same office in l993 as at the time of the survey, they were coded a missing value for this variable, to reduce risk of endogeneity. (I do not, however, have data about changes in workgroup within an office over the period of time, so some risk remains.)

24. The procurement workforce had increased from about l6,900 to about 27,400 between l980 and its peak in l989, and then declined by almost 2000 by September l993, just as the first reinventing government report was being issued. I would like to thank Linda Neilson of the Defense Acquisition University for providing these numbers. The numbers are for fiscal years, which end on September 30, which means l993 numbers are as of September 30, l993.

26. As in Chapter Four, the variable was recoded to go back to the earlier period. The slight change was that the cutoff year for working in one's current buying office was l994, rather than l993, to reflect looking at the influence of the most-respected coworker when early experiences with reform were taking place, not at the time of initiation of reform.

27. This was for the local office head in place as of l994. Only respondents in their current organization since at least the year after the change effort began were assigned a value (others were missing values).

28. I made the same change as for most-respected coworker attitude, including only those who were in their current workgroups as of l994.

29. An analogous recalculation was made.

33. In some cases, the judgment of whether a trait can discourage as well as encourage successful experience depends on conclusions about an appropriate "no effect"/neutral baseline -- what would happen absent presence of the trait being measured – that are not obvious. Probably the most appropriate "no effect"/neutral baseline for one's education or job level in the organization is the lowest-possible level (for example, no education at all), from which any increment would be hypothesized to promote successful experience. But, if one assumed a different neutral baseline, a negative effect would become possible, although in this example, the difference would simply result from a different assumption about a baseline and wouldn't (as in the case of opposite effects of the self-fulfilling prophecy) actually reflect different operation of a psychological mechanism. The judgment of whether a trait can discourage as well as encourage successful experience may also depend on one's interpretation of the wording of the question in the survey. Take being driven to succeed, where the survey question is "I am very strongly driven to succeed at my job." What can we say about a respondent who strongly disagrees with that statement? Is a "neutral" baseline being somewhat driven to succeed (which would roughly be equivalent to a "3" answer) or not being driven at all (a "5")? If the baseline is a "5," then even expressing mixed feelings regarding the statement might promote positive experience, while disagreement would not make one's experience

worse than "it otherwise would have been" if the trait were exerting no effect. But if neutral is being "somewhat driven to succeed," the middle response should serve as a baseline, as it does with the variable measuring self-confidence.

If the regression model shows, say, a positive coefficient for a variable, this tells us that the more the characteristic is present, the more successful the experience. But this does not suggest that at all values of the characteristic, experience is better than it "otherwise would have been," since the scaling used for these variables is arbitrary. (If the variable had been scaled the opposite way, the coefficient would have been negative, but that wouldn't have meant that at all values of the characteristic, the value of the dependent variable was lower than otherwise.) Instead, we must think deductively about what is plausible. I discussed earlier why one would expect that, for some, operation of the self-fulfilling prophecy would produce worse, not better, experiences, while we would hypothesize supervisor support to promote better experience but supervisor opposition not to promote a worse one. For variables we conclude can exert both positive and negative effects on experience, we can specify using logic a "no effect"/neutral value for the variable, such as self-confidence=3, reflecting a respondent neither self-confident nor unconfident. If there is a significant coefficient and the variable is operating symmetrically, then if the mean sample value for the variable is below 3.00 (reflecting, given the question wording, greater self-confidence), this trait would on balance be promoting successful experience, above 3.00 on balance promoting unsuccessful experience.

In addition to curves where the relationship between x and y have the same sign at all values of x, though a different slope, it is also possible for the polynomial to define a U-shaped or upside-down U-shaped curve, where the sign of the relationship between x and y changes at different values of x.

34. To determine the net impact of the variable on y, one simply adds up the impact on y for each respondent, given the polynomial coefficients, and divides by the number of respondents. For the reasons noted above, one must rescale the variable so that the neutral value is zero in order to be able to compare the result with the hypothetical linear function with a neutral effect. This allows a statement about the impact of the variable in question in the actual data being examined, but not any generalized statement about whether, at some lower mean value, the impact of the variable on y would necessarily have been positive or negative.

It would be nice to be able to calculate a value for what the sample mean value of the predictor variable would need to be, above which the overall effect of the variable on the dependent variable became positive -- so that one might say, for some variable with an asymmetric impact, that if the variable had a sample mean greater than 2.5 (on a l-5 scale), it was on balance increasing the value of the dependent variable, compared to a situation where the variable behaved symmetrically and had a mean value with a neutral effect. (Actually, to make this comparison, it would be necessary to rescale the variable in question so that its neutral value is zero – e.g. rescale a variable to which the values l-5, with 3 as a neutral value, have arbitrarily been attached, to a scale from –2 to +2, with 0 as the neutral value. The reason is that, though we know the coefficients of an actual nonlinear model, based on the actual data being analyzed, with coefficients attached to the x^2 and x terms of the variable being examined, there are an infinite number of hypothetical linear models, with different coefficients attached to the x term of the variable being examined, that could correspond to any one of an infinite number of hypothetical linear relationships between that x variable and the dependent variable. The only way that we can make a comparison to "a" hypothetical linear relationship is if the variable in the linear model takes the mean value zero, in which case, no matter what the coefficient, the net impact of that variable on y will be zero, i.e. neutral.)

This is, however, not possible, because, where the relationship between x and y is nonlinear, a large range of mean values for x will all be consistent with a neutral effect on the dependent variable. Given the presence of a squared term, different distributions of values for x will predict different values of y, even if those distributions have the same mean value, because (for a variable that takes only positive values) distributions with many high values for x and low values for x will produce, because those high values are squared, a higher predicted value for y, compared to a distribution with the same mean but with more values clustered around the middle. This can be seen by considering a simple example. Take a sample of three observations where the three values x takes are 2, 2, and 5. This sample will have a mean of 3; the squared terms for these three values (ignoring the linear term) will add up to 33 (4 plus 4 plus 25), giving a mean value of 11 for y. A sample where the three values x takes are 3, 3, and 3 will also have a mean value of 3; here, however, the squared terms will add up to 27(9 plus 9 plus 9), giving a mean value of 9 for y. Given the range of predicted values for y consistent with any given mean value for x, it will also be the case that, if one takes a continuum of adjacent possible mean values for x, each with an associated range of values for y, overlap of the y's will occur -- there will be more than one mean value for x associated with a given predicted mean value for y. I am grateful to my colleague Chris Avery and my research assistant Chris Hans for helpful discussions and clarifications of this issue.

Using a mathematical theorem called Jensen's inequality, one may establish the highest sample mean value of x that would be consistent with a neutral effect, such that at any higher mean value, the net impact of x on the dependent variable would have to be positive. Likewise, one may, using the same theorem, establish the lowest sample mean value of x that would be consistent with a neutral effect of the predictor variable on the dependent variable, such that at any lower mean value, the net impact of x on the dependent variable would have to be negative. (In other words, this is the range of possible values below which the change in y, given the distribution of x, can't be positive and above which the change in y must be positive.)

According to Jensen's inequality, for a curvilinear function that takes a concave form (which the ones we will be discussing, except for the one involving initial experience in the empowerment model, all will turn out to do, as we shall see later), the predicted (expected) value of the function, at any constant mean value, will be at a minimum where the only values of x that occur are the extreme positive and the extreme negative values x can take; it will be at a maximum where the neutral value zero occurs with certainty (i.e. for all observations in the sample). Since it takes a distribution of x values consisting only of the extreme values to produce the minimum expected value associated with a given mean, it follows logically that, if you have a distribution including less extreme values of x, in order to get the same (low) expected value of x, the mean value of x would need to be lower than the original given mean. Therefore, Jensen's inequality also implies that highest mean associated with a given expected value – which, setting the expected value at zero, the neutral effect, is the quantity for which we wish to solve – will be the mean associated with a sample that contains only the two most extreme distributions. (The opposite is the case for obtaining the lowest mean associated with the given expected value.)

To solve for the maximum mean value consistent with an expected value of the neutral effect of zero, it is most straightforward to rescale the variable, so that the neutral value becomes "0" and other values are rescaled accordingly. (Thus, for example, the variable PR, which was assigned a scale from 0 to 100, with 50 as neutral, gets rescaled so that 0=-50, 50=0, and 100=+50.) The expected value of a sample distribution involving variable x comes from adding each value the sample takes times the coefficients associated with x times the probability that the value appears in the sample. In a sample that only takes two values, this would reduce to $p(ax1^2 + bx1) + (1-p)(ax2^2 + bx2)$, where $x1$ and $x2$ are the two values the variable can take, and a and b are the coefficients of the polynomial function in the model.. Setting this expression equal to zero establishes an equation for the distribution of variables that are consistent with an expected value of zero. Since we see from the above that the highest mean value consistent with a given expected value, for a concave function, occurs where the only two values are the lowest and highest values the sample takes, we plug in the minimum and maximum values for the variable (in the case of a rescaled reform attitude variable, -50 and +50) and use the actual coefficients in the model in question. One then solves for p. Using the probabilities that the variable will take the maximum and minimum values, one can then determine (using the formula for the mean value, $px1 + px2$) the highest possible mean value of the function, given those probabilities, for the neutral mean value of zero. (For the lowest mean value consistent with a neutral effect, set all values in the sample at zero, and, since all terms then become equal to zero, the corresponding mean value itself becomes zero.) I am grateful to my colleague Chris Avery for helping me with this problem.

The range in-between these two values defines the set of mean values that might have a positive or a negative net effect on the dependent variable, depending on the actual distribution of individual responses associated with the mean value in question (a sample distribution where half the sample had the value –3 and the other half +5 might have a positive net impact on the dependent variable, while one where half the sample had –1 and the other half +3 might have a negative net impact, though both distributions have the same mean value). The problem is that this range of mean values that might in principle have either a positive or a negative net impact on the dependent variable can be expected to be so wide that it's not very helpful, because one could only be certain that the variable being examined was having a positive impact on the dependent variable at a very high mean value for that variable. For example, using Jensen's inequality, the possible range of values consistent with a neutral effect for the variable distortion (which, as we shall see, is one of the variables displaying an asymmetric effect) turns out to be 0 through 44.7, for a variable for which the positive values range between 0 and l00.

36. Also, in addition to testing for impact of demographic factors in creating positive feedback promoting good experience trying reform, education and job level also served as controls for an alternate explanation to operation of a self-fulfilling prophecy for any impact of initial attitude on initial experience, since education and job level, were correlated with initial attitude and might be correlated with initial experience.

38. It is possible that some individuals may have been sent on training oriented towards reform, or had volunteered for such training, because they supported reform, so endogeneity might have existed -- any causal arrow partly went from support for reform to receiving training, rather than the other way around. I therefore also tested an alternate specification of this variable, which left out attendance at off-site courses and use of the Internet for procurement reform training, on the grounds that (as opposed to staff meetings or on-site courses, which everybody would attend) more choice may have been involved in an individual's attendance. This created a scale that took values from zero to 6. For all the three models where results will be shown here, results were very similar to those that will be reported for the specification of this variable used here, though p-values were sometimes less strong.

Local office head attitude and supervisor attitude also controlled for the possibility pro-reform leaders organized more training and that having a pro-reform leader rather than receiving the training was determining the results.

39. Since the goal was to see whether any kind of training made a difference for respondent attitudes and/or behavior regarding reform, the decision was made to examine mean values on the reform attitude and reform behavior variables for respondents who had received or not received each kind of training, so as to select from the various forms of training those that, at least on a bivariate basis, had the largest impact on these variables. These were off-site course, courses given at work, brown-bag lunches/staff meetings/informally from colleagues at the office, and the Internet/CD-ROM's. (Off-site conferences and satellite broadcasts/videotapes were excluded.)

The coding order used was created to correspond to linear changes seen in the mean values in the exploratory data analysis, where those receiving only reform training had somewhat higher scores for these variables than those who had received both kinds of training.

The construction of this variable makes it more an ordinal than an interval scale. (An interval scale is one where distances between points on the scale are equal, normally required in the social sciences for confident conclusions from a regression analysis. An ordinal scale is one where a "2" has a value clearly lower than a "l," but distances between the points on the scale shouldn't be assumed to be equal.) This, as well as the lack of more specific information about the nature of training, suggests caution interpreting coefficients for this variable.

40. In the overall experience model, supervisor attitude measured the self-reported attitude of one's current supervisor as of the time of the survey, and the recoding of most-respected co-worker attitude used in the initial experience model was eliminated.

71. Granovetter (l978) introduced the idea that different people have different “thresholds” for participating in a new activity. He defines (Granovetter 1978: 1422) an individual’s threshold as “the proportion of the group he would have to see join before he would do so.” A person’s threshold is based on some mixture of the person’s underlying attitude towards the change and his or her degree of venturesomeness or risk aversion. So, to cite one of Granovetter’s examples, some will be willing to join a anti- dictatorship and also less risk-averse) while others will require many to join (they are moderately anti-dictatorship and also more risk-averse). People with low

thresholds may be seen either as only noticing anti-dictatorship behaviors (and ignoring the many who do nothing) or as noticing both, but being willing to act with a smaller proportion of others acting than would more risk-averse people.