McGregor's Theory Y vs. Bentham's Panopticism: Toward a Critique of the Economic Theory of Agency

David Ellerman

Introduction

This paper is part of a larger project to better understand the limitations of the economic theory of agency and incentives.[1] The economic approach focuses on extrinsic incentives whereas a better understanding of human organization requires an understanding of intrinsic motivation and the complementary or substitutive relationships with extrinsic motivation. I will focus on different treatments of informational or "panoptic" questions regarding transparency or non-transparency in the management of firms or other human organizations. This context will give the questions definiteness but the ideas might also be applied in a broader social context.

As a conceptual framework, I will use the seminal work of Douglas McGregor. The economic approach to agency and incentives theory is based essentially on what McGregor called the "Theory X" worldview, and the alternative explored here is essentially what he called the "Theory Y" view of individuals in organizations.

The Economic Theory of Agency and Incentives

Agency theory focuses on the common situation wherein one person or group, called the "principal," desires to obtain certain behavior from another person or group called the "agent." The principal-agent language is borrowed from the legal relationship of agency and is used in economics

---

David Ellerman attended MIT and Boston University, earning masters in philosophy and economics and a doctorate in mathematics. He is Economic Advisor to the Chief Economist of the World Bank and the author of *Intellectual Trespassing as a Way of Life: Essays in Philosophy, Economics, and Mathematics* (Lanham, MD: Rowman & Littlefield, 1994) and many other books.

Knowledge, Technology, & Policy, Spring 2001, Vol. 14, No. 1, pp. 34-49.

in a much broader context.[2] McGregor, writing in 1948 before the principal-agent language was established in the economics and management literature, refers to the principal and agent respectively as "*A*" and "*B*": "*A* always refers to the individual (or group) who is attempting to induce a behavior change, and *B* always refers to the individual (or group) whose behavior is affected." [1948; reprinted in 1966, 155]

Agency theory is based on *homo economicus*, or in McGregor's terms, on the Theory X view of people. The individual agents are assumed to dislike work so positive and negatives incentives ("carrots and sticks") must be supplied by the principal to induce the appropriate behavior by the agents. Left to their own devices, agents cannot be trusted to act in the manner desired by the principal so an incentive structure must be applied to redirect the agents in the desired manner.

Information enters the picture because the principal needs to know if the agents are carrying out the desired actions or at least are delivering the desired outcomes or results. The principal's information about the agents is always in fact incomplete so incentive structures must be designed with this information asymmetry in mind. There are two broad types of informational deficiencies. When the principal lacks information about something that is within the discretionary choices of the agent, then it is broadly called a "hidden action" or "moral hazard" situation. If the principal lacks information about something that the agent cannot change by taking different actions, then it is broadly called a "hidden characteristic" or "adverse selection" situation.[3] The concern here is with hidden action problems where, within the penumbra of the principal's ignorance, the agent or agents might act opportunistically to the detriment of the principal's interests.

The economic theory of agency works with variables such as monetary rewards that the principal can affect and change since the goal of the theory is to design incentive structures to elicit the desired results from the agents. By following the incentives provided by the "incentive-compatible" reward scheme, the agents will be led to achieve the results desired by the principal. From the agents' viewpoint, such an incentive structure represents extrinsic or external motivation.

> Money is the most obvious [extrinsic motivator] but promotion, praise, recognition, criticism, social acceptance and rejection, and 'fringe benefits' are other examples.
>
> 'Intrinsic' rewards, on the other hand, are inherent in the activity itself; the reward is the achievement. They cannot be *directly* controlled externally, although characteristics of the environment can enhance or limit the individual's opportunities to obtain them. Thus, achievement of knowledge or skill, of autonomy, of self-respect, of solutions to problems are examples. [McGregor 1966, 203-4]

There is now a considerable body of literature in psychology, sociology, and organizational behavior on intrinsic and extrinsic motivation as well as the closely related notions of autonomy, self-determination, and internal locus of control.[4] Although considerations of intrinsic motivation have figured prominently in the Romantic critique[5] of classical economics, the topic has until recently only received sporadic treatment in economics literature.[6] Bruno Frey's recent *Not Just for the Money* [1997] is the first book-length treatment of the topic of intrinsic motivation in the economics literature.

The important point to notice is that incentive structures devised by the principal to control the agents must rely, by necessity, on extrinsic motivation. Any actions by the agents that were *motivated* by the incentive structure provided by the principal would, *ipso facto*, not be activities done for their *own* sake—as their own reward. Let me repeat this crucial point; actions done for the sake of incentives provided by the principal are not action done for their own sake. That is the crux of the inherent limitation of agency theory; external control and intrinsic motivation (which requires an internal locus of control) do not mix.

External Control and Panopticism

External control by the principal requires information obtained by monitoring the agents, and that, in turn, requires a certain transparency on the part of the agents. Jeremy Bentham was perhaps the strongest advocate of publicity and transparency: "without publicity, no good is permanent; under the auspices of publicity, no evil can continue."[7] His Panopticon [1995 (1787)] prison scheme (cells arranged around the rim of a circular building all visible from a watchtower in the center) represents the apex of this theme in his writings and thus the emphasis on the one-way monitoring and supervision of a controlled group of people (e.g., agents) has been called "Panopticism."[8] Bentham's plan is often thought to be only an eccentric idea for a prison but he was quite explicit about more general applications.[9] Indeed, the long title of the work begins "Panopticon; or, The Inspection House: containing the Idea of a New Principle of Construction Applicable to any sort of Establishment, in which Persons of any Description are to be kept under Inspection; ..." [1995 (1787)] and then he goes on to list "Houses of Industry" "Manufactories," "Hospitals, and Schools" in addition of "Penitentiary-Houses" and "Prisons." The panoptic principle is "a great and new instrument of government...; its great excellence consist in the great strength it is capable of giving to *any* institution it may be thought proper to apply it to." [Bentham quoted in Foucault 1977, 206]

In a factory, the requirements for monitoring would depend on the nature of the incentive scheme.

> Whatever be the manufacture, the utility of the [inspection] principle is obvious and incontestible, in all cases where the workmen are paid according to their *time*. Where they are paid by the *piece*, there the interest which the workman has in the value of his work supersedes the use of coercion, and of every expedient calculated to give force to it. In this case I see no other use to be made of the inspection principle, than in as far as instruction may be wanted, or in the view of preventing any waste or other damage, which would not of itself come

> home to the workman, in the way of diminishing his earnings, or in any other shape. [Bentham 1995 (1787), 80-1]

Where monitoring and detailed control is quite costly or not feasible (and where the workman will not pay a franchise fee to be an independent franchisee, e.g., taxicabs), then some form of piece rate or performance-related pay is recommended by agency theory.[10] Since the effort expended by the agent is hidden,"any measure of performance that (on the margin) reveals information on the effort level chosen by the agent should be included in the compensation contract." [Prendergast 1999, 12-13] In that manner, the economic results of the effort decisions would"come home to the workman."

In a survey article, Eisenhardt [1989] notes that, in addition to emphasizing considerations of risk, agency theory emphasizes the role of information.

> This gives an important role to formal information systems, such as budgeting, MBO[11], and boards of directors, and informal ones, such as managerial supervision, which is unique in organizational research. The implication is that organizations can invest in information systems in order to control agent opportunism. [64]

Information about the actions of the agents allows the principal to more closely specify job performance, to design higher powered incentives, and to better curtail opportunistic behavior. Since information is costly, the principal does not seek full "panoptic" knowledge. Monitoring would in general be pushed to the point where the marginal cost of information equals the marginal return in better control of agent behavior.[12]

Agency theory falls short of recommending full panoptic transparency only because of"agency costs"such as the costs of monitoring and the costs of writing and enforcing more complete contracts. As improving information technologies reduce these agency costs, the recommendations of the economic theory of agency and incentives will more closely approximate the panoptic ideal of full monitoring and control of the agents by the principal.

Intrinsic Motivation and Theory Y

An intrinsically motivated activity is an activity carried out by individuals for its own sake. The activity is an end in itself, not an instrumental means to some other end (such as satisfying biological needs or "tissue deficits"). The factors that determine the meaning of"for its own sake"are usually based on the self-identity of the person or persons carrying out the activity. An intrinsically motivated activity might be accompanied by extrinsic motivators if the latter are not controlling, i.e., if they do not take over the locus of control. For instance, professors typically pursue their professional work for its own sake even though they also receive a salary and other emoluments. Indeed much of the story is concerned with the question of the locus of control for an activity. Autonomous activity has an internal locus of control. A bribe to do what one would not ordinarily do establishes an external locus of control.

The roots of intrinsic motivation such as an individual's self-identity (including the larger social units with which the person identifies) are typically not open to intentional and deliberate choice. One chooses according to who one is, but one does not directly choose who one is. These basic "preferences" can be transformed but more as a "by-product of actions undertaken for other ends" [Elster 1983, 43] than as the result of deliberate actions.[13] For instance, one cannot simply decide to be "in love" and thus one cannot "buy love." This "can't-buy-love" situation limits the domain of the market and the reach of extrinsic motivators. "Carrots and sticks" might buy or induce compliant (e.g., "loving") behaviors, but they cannot supply or create intrinsic motivation–although we will see in the next section how "carrots and sticks" can override intrinsic motivation and may eventually cause it to atrophy.

In the applications of agency theory to the workplace, the hidden actions of the workers are typically concerned with the effort level of the work. For instance, there might be quality bonuses that would reward employees for producing at a certain level of monitored quality in the products. Yet what are loosely called "Japanese-style" management techniques (e.g., often associated with W. Edwards Deming) go in quite the opposite direction and achieve world-renowned levels of quality. For instance, Deming's "New Economics" recommends to "Abolish incentive pay and pay based on performance" [1994, 28], e.g., to pay salespeople by salary rather than by commission. Deming recommends replacing a system based on monitoring and quality bonuses with a system using (for the most part) trust based on self-esteem and pride in the quality of one's work. In short, this approach to quality relies not on cleverly constructed pay-for-performance schedules but on switching over to a quality system driven largely by intrinsic motivators such as self-esteem and pride in one's work–in short on quality as a calling.[14]

While Japanese-style management methods and Deming's work has done much to popularize the role of intrinsic motivation in management and to criticize the pay-for-performance schemes, the most sustained development of these themes is McGregor's Theory Y [1948, 1960, 1966, 1967].[15] It is not a matter of choosing extrinsic or intrinsic motivation; it is a matter of background and foreground. When basic physiological and security needs are not satisfied, then such motivations will stay in the foreground. "Man tends to live for bread alone when there is little bread." [1960, 41] But when these needs lower on the Maslovian hierarchy are adequately and equitably satisfied, then the higher needs" of self-actualization and self-fulfillment move into the foreground.[16] And when extrinsic motivators such as pay are in the background, then the perception of equity may be more important than any purported linkage with performance. Indeed, the detailed linkage of pay to performance, as per agency theory, is pointless when pay is in the background and performance is based on intrinsic motivation.

> [T]he carrot and stick theory does not work at all once man has reached an adequate subsistence level and is motivated primarily by higher needs. Management cannot provide a man with self-respect, or with the respect of his fellows, or with the satisfaction of needs for self-fulfillment. It can create conditions such that he is encouraged and enabled to seek such satisfactions *for himself*, or it can thwart him by failing to create those conditions. [McGregor 1960, 41; 1966, 13]

That is why agency theory is not just a harmless "academic toy." It makes misleading recommendations both concerning the pay-performance linkage and the desirability of individual monitoring of agents by principals. The point is not an improved behaviorism or Theory X approach promised by agency theory, but a wholly different approach along Theory Y lines.

Since the satisfaction of intrinsic motivators is inherently a do-it-yourself project, the principal cannot motivate, direct, or control the agents in those activities. Management has an indirect role:

> The essential task of management is to arrange organizational conditions and methods of operation so that people can achieve their own goals *best* by directing *their own* efforts toward organizational objectives.
>
> This is a process primarily of creating opportunities, releasing potential, removing obstacles, encouraging growth, providing guidance. [McGregor 1966, 15]

A person's intrinsic motivation is based on the person's own definition of self-identity. Yet through identification, a person's definition of self-identity can extend to broader groups such as family, ethnic or professional group, and organization (e.g., being a "Toyota man" or a "neoclassical economist"). Herbert Simon [1991] has developed a scathing analysis of the explanatory effectiveness and organizational practicality of agency theory and the related new institutional economics.

> Although economic rewards play an important part in securing adherence to organizational goals and management authority, they are limited in their effectiveness. Organizations would be far less effective systems than they actually are if such rewards were the only means, or even the principal means, of motivation available. In fact, observation of behavior in organizations reveals other powerful motivations that induce employees to accept organizational goals and authority as bases for their actions. [The] most important of these mechanisms ...[is] organizational identification. [Simon 1991, 34]

He goes on to note specifically that the "attempts of the new institutional economics to explain organizational behavior solely in terms of agency, asymmetric information, transaction costs, opportunism, and other concepts drawn from neoclassical economics ignore key organizational mechanisms like authority, identification, and coordination, and hence are seriously incomplete." [Simon 1991, 42]

Potential Negative Effects of Management by Carrots and Sticks

Agency theory's neglect of intrinsic motivation would be of less concern if extrinsic and intrinsic motivation were always additive or complementary. While they may be complementary when extrinsic motivators are largely satisfied and in the background, negative effects arise when the extrinsic motivators are brought to center stage as instruments of control.

While management cannot provide intrinsic motivation, it can nevertheless frustrate the satisfaction of those drives by imposing a management structure based on direction and control, i.e., Theory X (and agency theory).[17]

> The philosophy of management by direction and control—*regardless of whether it is hard or soft*—is inadequate to motivate, because the human needs on which this approach relies are relatively unimportant motivators of behavior in our society today....
>
> People, deprived of opportunities to satisfy at work the needs that are now important to them, behave exactly as we might predict—with indolence, passivity, unwillingness to accept responsibility, resistance to change, willingness to follow the demagogue, unreasonable demands for economic benefits. It would seem that we may be caught in a web of our own weaving. [McGregor 1960, 42]

It is not a question of whether Theory X or Theory Y is true. Both theories might be taken as rough descriptions of two different organizational equilibria. Theory X is a low trust and low involvement equilibrium, and Theory Y is a high trust/involvement equilibrium. Distrust breeds distrust and trust breeds trust so both equilibria are self-reinforcing. When McGregor indicates that "we may be caught in a web of our own weaving" he means that when we assume people behave as indicated by Theory X (e.g., distrustful) and thus impose a Theory X managerial structure, we may soon induce the assumed behavior even if it was not initially present. Distrust breeds distrust in the Theory X vicious circle.

> The introduction of pay into a task situation, then, moves, or seems to move, the locus of causality from disposition to circumstance, from internal to external; it alters the task from chosen to unchosen, and since people do not work without motives, it transmutes intrinsic motivation into pecuniary motivation. Strangely, it *creates* Skinnerian man where he was missing earlier. [Lane 1991, 379]

Thus agency theory and Theory X management can usually find suitable empirical (pseudo-) verification for the *homo economicus* presuppositions.

Money talks—and when it talks loud enough and long enough, people will be able to hear little else.

Let us consider more closely how the use of extrinsic motivation for control purposes drives or crowds out intrinsic motivation. The notions of internal or external loci of control (or causality[18]) are related to intrinsic and extrinsic motivation. When an externally offered reward or punishment is used successfully to redirect a person's behavior, then the behavior is said to have an "external locus of control." When one acts for one's own reasons and is the source of one's actions, then one would be said to have an "internal locus of control." Having an internal locus of control is to act autonomously in contrast to responding to heteronomously imposed rewards or punishments.[19]

External interventions by other people intended to change a person's behavior pose a threat to autonomy. The threat-to-autonomy or reactance [see Brehm 1972] effect results from using extrinsic motivators—carrots and sticks—to shift the locus of control from internal to external.[20] The effect shows itself in a poor quality and low effort performance, in sullen and perfunctory behavior fulfilling the letter but not the spirit of an agreement, and perhaps even in the urge to defiantly do the opposite just to show one's autonomy.

The threat-to-autonomy effect points to a broader complication in accounting for human preferences. An individual's evaluation of an event may be strongly affected by the source of the event (e.g., was a death due to natural causes, an accident, or a murder). A change in a person's choices due to another person's strategic action may give rise to a "reactance" that would be absent if the change had been necessitated by natural events.

The imposition of extrinsic motivation may have untoward long-term "by-product" effects. Intrinsic and extrinsic motivation might not be additive. Indeed it frequently seems to be the case that extrinsic incentives superimposed onto a system involving intrinsic motivation in order to better achieve control will tend to "crowd out"[21] and atrophy the intrinsic motivation.

> It was on the basis of this atrophy dynamic—the less the requirements of the social order for the public spirit, the more the supply of public spirit dries up—that the United States' system for obtaining an adequate supply of human blood for medical purposes, with its only partial reliance on voluntary giving, was criticized by the British sociologist Richard Titmuss. And the British political economist Fred Hirsch generalized the point: once a social system, such as capitalism, convinces everyone that it can dispense with morality and public spirit, the universal pursuit of self-interest being all that is needed for satisfactory performance, the system will undermine its own viability, which is in fact premised on civic behavior and on the respect of certain moral norms to a far greater extent than capitalism's official ideology avows. [Hirschman 1992, 155-6]

This crowding-out or atrophy effect might also be amplified by what Frey calls a "motivational spillover effect." The imposition of controlling extrinsic incentives might not only atrophy intrinsic motivators in the given system but in related areas of endeavor where the market-type incentives were not applied.

In summary, the principal cannot directly supply or increase the intrinsic motivation of the agent, but the principal can reduce intrinsic motivation (e.g., displace it into the motivational background where it may atrophy) or prevent it from moving into the foreground. Intrinsic motivation can be crudely displaced by threats or subtly displaced by the offer of what economists tellingly call "high powered" incentives (e.g., stock options to focus managers on the stock market and to weaken any "conflicting" ties with other stakeholders such as the enterprise personnel, the customers, the suppliers, and the local community).[22]

The Theory Y Case Against Panopticism

We are now is a position to see that Theory Y has quite different implications for "panoptic" monitoring of the agents than the Theory X approach of the economic theory of agency and incentives. An agency theoretic management system based on monitoring and control may start the dynamics that causes a Theory Y high-trust virtuous circle to decay into a low-trust Theory X vicious circle.

> Intrinsic rewards are largely ignored under these circumstances. Surveillance displaces autonomy, mistrust undermines self-regard, absence of support and help minimizes achievement, likelihood of punishment for noncompliance reduces risk-taking and innovation, rigidity of standards and administrative procedures precludes the individual's use of his own know-how. [McGregor 1967, 126-7]

The close monitoring and supervision of the agent will reveal distrust by the principal and may lead to a negative reaction on the part of the agents and eventually to an atrophy of any intrinsic motivation for the activities.

McGregor develops this argument in the context of his analysis of the staff-line relationship. [1960, chapters 11 and 12] A staff member, the helper, is called upon to help a line manager, the doer, to address and solve an organizational problem. The line manager also has a manager and the question is how "transparent" should the doer's (line manager's) activities be to the manager. In the course of working with the line manager, the staff member (helper) will compile information and may even set up information monitoring systems (e.g., accounting or finance information) about the line manager's activities. Who should receive this information?

Theory Y is based on the idea of maximizing each person's autonomy and self-control within the organizational framework.

> With respect to data and reports compiled by staff groups, the principle of self-control requires that they be provided to each member of management for controlling *his own*, not his subordinates' jobs. [McGregor 1960, 160]

Thus the informational feedback about a person's performance should go to that person for the purposes of self-control, not to the person's superior. The system has a built-in "blindness" or non-transparency. In that manner, the informational feedback is enabling to the subordinates rather than controlling.[23] The information that goes back to a division manager should be aggregate, not individuated.

> If such summary data indicate to the manager that something is wrong within the organizational unit for which he is responsible, he will turn not to staff [i.e., not to the accounting or finance staff providing the data], but to his subordinates for help in analyzing the problem and correcting it. He will not assign staff "policemen" the task of locating the "culprit." If his subordinates have data for controlling their own jobs, the likelihood is that they will already have spotted and either corrected the difficulty themselves or sought help in doing so. [McGregor 1960, 161]

McGregor laments that that there is so much managerial misuse of information by line or staff managers who seem not to understand the non-transparency implications of the principle of self-control or not to be committed to the principle. Managers complain that "delegation" doesn't seem to work—that subordinates "don't want responsibility"—when the same managers maintain surveillance through various performance measures of their subordinates.[24] One of the prerequisites for a person to take responsibility for the job is for the person, not the person's superior, to have necessary information to control the job.

The members of the staff departments should provide information only for the purposes of self-control, not the control of others.

> The same principle—that staff provides help for self-control only—applies to what is usually called "coordination," but which means policing the organization with respect to policy and procedures. Help can consist in informing an individual that he is out of line, or that a contemplated action would be in violation of policy—*but with the full understanding by both parties that the staff member will not report his knowledge or opinion to anyone else.* [McGregor 1960, 169]

The helper cannot at the same time be a policeman exercising control as that frustrates, rather than facilitates, the self-control of the doers.

The Panoptic Power of Information Technology

These questions of where there should be transparency in an organization—and where there should not be—will only grow more acute in the future. With the development of modern information technologies, there is the potential both for better informational feedback to enable self-control or for even more abuse by managers of the principle of self-control. In a chapter entitled "The Information Panopticon" of her book *In the Age of the Smart Ma-*

chine, Shoshana Zuboff notes that information technology codifies the events and processes of work as a "text" that managers and workers can read.

> To what extent was it the occasion for organizational members to more fully grasp their own reflections, learn from what they could now see, and increase their opportunities for autonomous actions? To what extent would the text be treated as a technical convenience enabling more assiduous behavior control? Was the text to be the keystone of a new learning environment, or would the increased visibility of behavior be exploited in the service of managerial control as an antidote to the pressures of uncertainty? [Zuboff 1988, 319]

She makes the distinction between information technology that is used to *automate*—increasing the panoptic power and control in the hands of managers—or to "*informate*"— increasing the self-control and learning opportunities of the operators carrying out the tasks. Managers who control how information technology will be used have shown a natural tendency to seek Bentham's "universal transparency"—at least insofar as their subordinates are concerned. Zuboff quotes an ecstatic manager:

> It is beautiful now. I can track my people's work. All I have to do is to type the [subordinate's] initials in and see how he is progressing and see what his total work load was. What is his productivity? Before, we had to judge people more on hearsay. Now we have it in black and white. [1988, 331]

But the managers have an equally natural reticence to have this information in the hands of *their* boss. Another manager notes:

> I don't think my boss should get involved with it....If anything is wrong, he can come and ask me. I don't think he should be involved with the nitty-gritty of running the job on my level. [340]

Thus the power and information dynamics of the new information panopticon are played out. Often any information that *can* be made technologically available to managers *is* made available to them in a renewed dream of panoptic power, a power based not on direct personal observation but on indirect technologically mediated surveillance. The violation of the principle of self-control inevitably brings resistance.

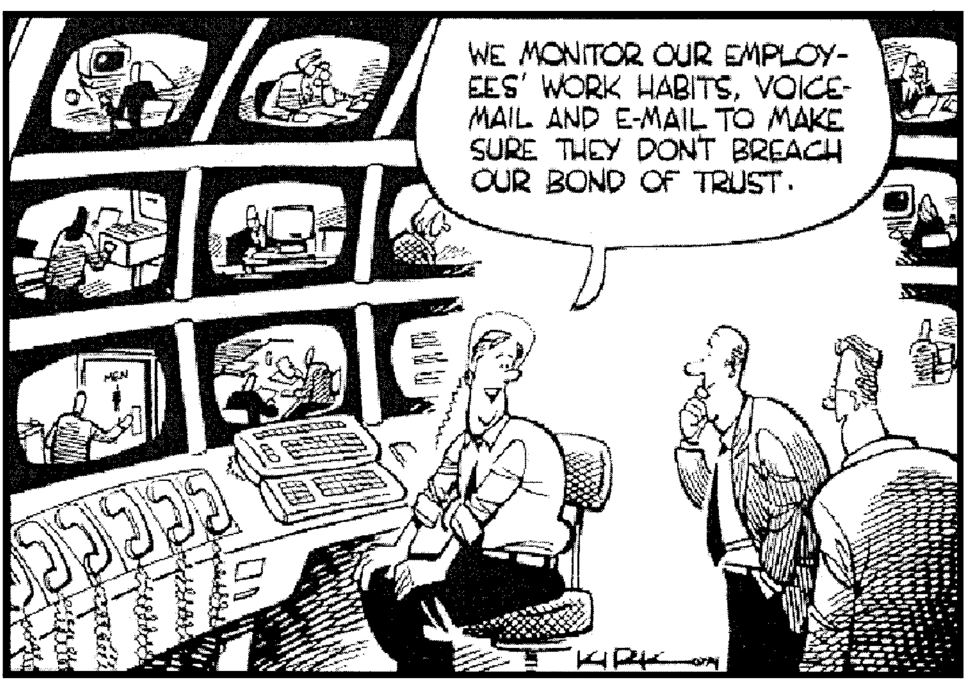
> The resistance to such exposure reflects in part an effort to retain a sense of self-control and to avoid feelings of shame....One way to minimize the risk of shame would be to look for ways to circumvent the observer, to thwart the power of the panopticon. This motivates managers who seek a technological solution that will ensure some measure of shadow, of privacy, and thus of self-control, in the increasing glare of the information panopticon. [Zuboff 1988, 344-5]

While some managers and workers might be able to informationally protect some sphere of autonomy, the whole "non-transparency" issue of managerial non-access to information about subordinates' performance has rarely been faced squarely.

> Not one of the organizations I studied had yet confronted this crucial issue. Rules of access were being constructed ad hoc, as individual managers were motivated to dip into data normally reserved for their subordinates and as subordinates found it in their interests to restrict or impede their superiors' access. Most people said that it would "take a lot of guts" for their organizations to develop a clear policy for data access. [Zuboff 1988, 357]

Conclusion

The idea of autonomy and self-control, encapsulated in McGregor's Theory Y, provides an argument for limiting the "transparency" of the agents activities to the principal—or rather for building a relationship where individuated monitoring of the agent by the principal is not part of the *modus operandi*. This non-transparency argument seems considerably more subtle and nuanced than the traditional arguments for privacy based on preventing inappropriate influence. The argument has nothing to do with the agent's private matters (i.e., outside the workplace). In this context, the principal (e.g., management) has an appropriate influence on the agents (e.g., workers)—albeit indirect by creating the conditions for the agents to



promote the goals of the principal (e.g., the organization) by their own intrinsically motivated activities.

The themes we have seen developed as "Theory Y" in organizational management are actually quite general. The implications for "Theory Y managers" can be generalized to democratic leaders and other holders of power in organizations where the autonomy of subordinates is to be nourished and supported rather than curtailed and overridden. This general case against Panopticism is part of the general case against the exclusively economic theory of agency and incentives. The "scientific" recommendations of the economic theory of agency are rather one-sided, and organizations that put such schemes in the motivational foreground (rather than the background) will tend not to promote human development and autonomy.

There is another threat that lies beyond the scope of this paper but which should be mentioned to maintain perspective. From ancient times, there has been a contest over the issue of the transparency or legibility of the people's affairs to the state. States need to "read" social realities in order to raise taxes, provisions, and troops. The "high modernist" schemes of technocratic social engineering continue in the panoptic tradition to design organizations and projects to maximize the legibility and controllability of the citizens' activities to the state.[25] The all-seeing Big Brother of the Orwellian vision of *Nineteen Eighty-Four* found its realization in the totalitarian regimes of the Left and Right in the 20th century.

The drive to externally control people's actions is not limited to attempts to perfect schemes of extrinsic motivation. Far more insidious are the social technologies, enhanced by the ICT revoluiton, to more directly mold the determinants of intrinsic motivation–to promote identification with Big Brother–as well as to limit and impair people's cognitive abilities so that decisions are made on the basis of a well-engineered "view" of the world.[26]

Notes

1. See Ross 1973 and Stiglitz 1974 for early work. For surveys and applications, see Pratt and Zeckhauser 1991, Eatwell et al. 1989, or the text Campbell 1995. For earlier critical analysis of agency theory, see Perrow 1972, Hirsh et al. 1987, Pfeffer 1994, and the references contained in Eisenhardt 1989.
2. In the legal relationship, the agent takes on a legal role to act in the interests of the principal, but economists now use the terminology in a broader context where the agent is not necessarily under any legal obligation to act in the interests of the principal.
3. Thus the economic theory of agency and incentives is an important topic in the economics of information, e.g., Eatwell et al. 1989.
4. See Deci and Ryan 1985, Elster 1983, Lane 1991, Candy 1991, Kohn 1993, and Deming 1994.
5. See, for example, Ruskin 1985 (1862). Lutz 1999 gives an integrated treatment of Sismondi, Carlyle, and Ruskin.
6. See Titmuss 1970, Arrow 1972, Scitovsky 1976, Hirsch 1976, Sen 1982, Schelling 1984, Akerlof 1984, Hirschman 1992, Kreps 1997, and Prendergast 1999.
7. From Bentham 1843; quoted in Bok 1982, 174.
8. See Foucault 1977. See also Gandy 1993 and the references cited therein.

9. Nor was it just a passing fancy."On the Panopticon, ..., Bentham...centered his attention for over 20 years [between the ages of 20 and 40]. To it he gave up his time, his more ambitious works, and his fortune...." [Everett 1966, 45-6]
10. Whether such compensation schemes are practical is another matter. Herbert Simon points out that "such reward systems are effective only to the extent that success can be attributed accurately to individual behaviors.... But of course, intense interdependence is precisely what makes it advantageous to organize people instead of depending wholly on market transactions. The measurement difficulties associated with tying rewards to contributions are not superficial, but arise from the very nature and rationale of organizations." [Simon 1991, 33]
11. See later remarks on the non-vulgarized interpretation of Drucker's "Management By Objectives" or MBO.
12. This might involve some subtle points. For instance, Aghion and Tirole [1997] develop a model where increasing monitoring effort by the principal leads to delimiting the agent's scope of action and thus "stifling the initiative" for hidden action. Where the unmonitored activity of the agent still has a net benefit for the principal, the marginal return to more monitoring would have to take account of that indirect cost of stifling the agent's initiative in the determination of the optimal level of monitoring.
13. In terms of the analytical tools of economics, feasible choices determine the variables in a person's utility function, but the shape of the function itself is not subject to direct choice. Yet preferences or utility functions do shift over time as a by-product of actions taken and other influences.
14. The importance of intrinsic motivation has always been evident in the professions where intelligence, creativity, diligence, and empathy are important. For instance, given the patient's very limited ability to monitor a doctor's actions and the knowledge asymmetry between doctor and patient, a medical system based solely on extrinsic motivation would be hardly workable. As long as doctors can "bury their mistakes," even the reputational mechanism is a poor substitute for the intrinsic motivators of empathy and professionalism (e.g., as expressed in the Hippocratic Oath).
15. Peter Drucker [1954] developed essentially the same "Theory Y" ideas in his "Management by Objectives (MBO)" (also called "management by objectives and self-control") approach as opposed to "management by control" (as noted in McGregor 1966, 15-16 and in Drucker 1973). But the MBO theory was so popularized (indeed, vulgarized) by Drucker and others apparently in order to reach a mass market that it is commonly interpreted to mean "management by results" in a manner quite along the lines of Theory X and agency theory. Hence we will rely more on McGregor's treatment of these ideas.
16. See Maslow 1954, 1968.
17. This can be illustrated using the ancient neo-Platonic metaphor of an internal "fountain" with its source in the human self. Another person can create conditions and remove obstacles so that the fountain can flow or can effectively shut off the fountain altogether, but the other person cannot supply the *pressure* (motivation) that causes the inner fountain to flow in the first place. Another useful metaphor is biological growth. One can create the conditions to help a plant grow or one can destroy a plant, but one cannot directly force or control the growth.
18. See Deci and Ryan 1985 for the notion of locus of causality. They differentiate it from the notion of locus of control [see Lefcourt 1976] as they interpret the latter as dealing with the outcomes rather than sources of action. Nevertheless, one may find the notion of "locus on control" often used to indicate the source of actions. We use the notions of having an internal locus of control, self-determination, and autonomy as being roughly synonymous.
19. We use autonomy or internal causation broadly to include not only one's "original" integrated sense of self but norms based on social interactions which are eventually integrated and internalized so following those norms would come to have an internal locus of causality. See Deci and Ryan 1985, Chapter 5.

20. For example, suppose one buys a townhouse in the middle of Winter and is looking forward to Spring to spruce up the poorly attended front yard. But the Townhouse Association Beautification Committee [a.k.a."Lawn Nazis"] arrives before Spring to inform the new owners that they must attend to the yard or face certain penalties. Instead of just thinking "Now we have two reasons to spruce up the lawn," the new owners might well resent the attempt to externalize their locus of control.
21. See Frey 1997 passim.
22. Historically, the "Great Transformation" [Polanyi 1944] from pre-market to market societies and the current process of globalization can be usefully viewed through the lens of using "high powered" market-based incentives to override older identifications.
23. See Adler and Borys 1996 for a development of this theme.
24. It should be carefully noted that the principle of self-control is not an argument that managers should "delegate" responsibility and then maintain impersonal surveillance of subordinates through, say, accounting measures–as opposed to maintaining direct personal surveillance. It is an argument against both personal and impersonal surveillance of individual subordinates. "Self-control" is not an euphemism for more subtle external control.
25. See Scott 1998 for an extensive development of this theme.
26. There is a large literature on this topic but I would particularly recommend the stream of thought that runs through Dewey [1939], Lindblom [1990], and Chomsky [1987].

References

Adler, Paul S. and Bryan Borys 1996. Two Types of Bureaucracy: Enabling and Coercive. *Administrative Science Quarterly.* 41: 61-89.

Aghion, Philippe and Jean Tirole 1997. Formal and Real Authority in Organizations. *Journal of Political Economy.* 105(1): 1-29.

Akerlof, George 1984. *An Economic Theorist's Book of Tales*. New York: Camb. Univ. Press.

Arrow, Kenneth 1972. Gifts and Exchanges. *Philosophy and Public Affairs.* 1: 343-62.

Bentham, Jeremy 1843. On Publicity. In *The Works of Jeremy Bentham*. John Bowring Ed. Edinburgh: W. Tait. 2: 310-17.

Bentham, Jeremy 1995 (1787). *The Panopticon Writings*. Miran Bozovic Ed. London: Verso.

Bok, Sissela 1982. *Secrets*. New York: Pantheon.

Brehm, Jack 1972. *Responses to the Loss of Freedom: A Theory of Psychological Reactance*. Morristown NJ: General Learning Press.

Campbell, Donald 1995. *Incentives: Motivation and the Economics of Information.* New York: Cambridge University Press.

Candy, Philip 1991. *Self-Direction for Lifelong Learning*. San Francisco: Jossey-Bass.

Chomsky, Noam 1987. *The Chomsky Reader*. James Peck Ed. New York: Pantheon.

Dahl, Robert 1985. *Preface to Economic Democracy*. Berkeley: University of California Press.

Deci, Edward and Richard Ryan 1985. *Intrinsic Motivation and Self-Determination in Human Behavior.* New York: Plenum Press.

Deming, W. Edwards 1994. *The New Economics for Industry, Government, Education.* Cambridge: MIT Center for Advanced Engineering.

Dewey, John 1939. *Freedom and Culture*. New York: Capricorn.

Drucker, Peter 1954. *The Practice of Management*. New York: Harper & Row.

Drucker, Peter 1973. *Management: Tasks, Responsibilities, Practices*. New York: Harper & Row.

Eatwell, John, Murray Milgate, et al., Eds. 1989. *The New Palgrave: Allocation, Information, and Markets*. New York: W. W. Norton.

Eisenhardt, Kathleen 1989. Agency Theory: An Assessment and Review. *Academy of Management Review.* 14(1): 57-74.

Elster, Jon 1983. *Sour Grapes: Studies in the Subversion of Rationality*. Cambridge: Cambridge University Press.

Everett, Charles 1966. *Jeremy Bentham*. New York: Dell.
Foucault, Michel 1977. *Discipline and Punish*. Alan Sheridan Trans. New York: Vintage.
Frey, Bruno 1997. *Not just for the money: An economic theory of personal motivation*. Cheltenham: Elgar.
Gandy Jr., Oscar 1993. *The Panoptic Sort: A Political Economy of Personal Information*. Boulder: Westview Press.
Handy, Charles 1993. *Understanding Organizations*. 4th ed. London: Penguin.
Hirsch, Fred 1976. *Social Limits to Growth*. Cambridge: Harvard.
Hirsch, P., S. Michaels, and R. Friedman 1987."Dirty hands"versus"clean models": Is sociology in danger of being seduced by economics? *Theory and Society*: 317-36.
Hirschman, Albert O. 1992. *Rival Views of Market Society*. Cambridge: Harvard University Press.
Kohn, Alfie 1993. *Punished by Rewards: The Trouble with Gold Stars, Incentive Plans, A's, Praise, and Other Bribes.* Boston: Houghton Mifflin.
Kreps, David 1997. Intrinsic Motivation and Extrinsic Incentives. *American Economic Review.* 87: 359-65.
Lane, Robert E. 1991. *The Market Experience.* New York: Cambridge University Press.
Lefcourt, Herbert 1976. *Locus of Control*. Hillsdale NJ: Erlbaum/Wiley.
Lindblom, Charles. 1990. *Inquiry and Change*. New Haven: Yale University Press.
Lutz, Mark 1999. *Economics for the Common Good*. London: Routledge.
Maslow, Abraham 1954. *Motivation and Personality*. New York: Harper.
Maslow, Abraham 1968. *Toward a Psychology of Being*. New York: Van Nostrand.
McGregor, Douglas 1948. The Staff Function in Human Relations. *Journal of Social Issues.* 4(3): 5-22. Reprinted in McGregor 1966, 145-71.
McGregor, Douglas 1960. *The Human Side of Enterprise*. New York: McGraw-Hill.
McGregor, Douglas 1966. *Leadership and Motivation: Essays of Douglas McGregor.* Cambridge: MIT Press.
McGregor, Douglas 1967. *The Professional Manager*. Caroline McGregor and Warren Bennis Eds. New York: McGraw-Hill.
Orwell, George 1949. *Nineteen Eighty-Four*. London: Martin Secker & Warburg.
Patterson, Orlando 1991. *Freedom Vol. I: Freedom in the Making of Western Culture*. New York: Basic Books.
Perrow, Charles 1972. *Complex Organizations: A Critical Essay*. New York: McGraw-Hill.
Pfeffer, Jeffrey 1994. *Competitive Advantage Through People*. Boston: Harvard Business School Press.
Polanyi, Karl 1944. *The Great Transformation*. New York: Rinehart.
Pratt, John W. and Richard J. Zeckhauser, Eds. 1991. *Principals and Agents: The Structure of Business*. Boston: Harvard Business School Press.
Prendergast, Canice 1999. The Provision of Incentives in Firms. *Journal of Economic Literature.* XXXVII(March): 7-63.
Ross, Stephen 1973. The economic theory of agency: The principal's problem. *American Economic Review.* 63 (May): 134-9.
Ruskin, John 1985 (1862). *Unto this Last*. London: Penguin.
Schelling, Thomas C. 1984. *Choice and Consequences: Perspectives of an Errant Economist*. Cambridge: Harvard.
Scitovsky, Tibor 1976. *The Joyless Economy: An Inquiry into Human Satisfaction and Consumer Dissatisfaction*. Oxford: Oxford University Press.
Scott, James C. 1998. *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. New Haven: Yale.
Sen, Amartya 1982. *Choice, Welfare and Measurement*. Oxford: Blackwell.
Simon, Herbert 1991. Organizations and Markets. *Journal of Economic Perspectives.* 5(2 Spring): 25-44.
Stiglitz, Joseph 1974. Incentives and risk sharing in sharecropping. *Review of Economic Studies.* 41(April): 219-255.
Titmuss, Richard 1970. *The Gift Relationship*. London: Allen and Unwin.