

## Enforcing Rules on Oneself

Thomas C. Schelling

*Journal of Law, Economics, & Organization*, Vol. 1, No. 2 (Autumn, 1985), 357-374.

Stable URL:

<http://links.jstor.org/sici?sici=8756-6222%28198523%291%3A2%3C357%3AER00%3E2.0.CO%3B2-N>

*Journal of Law, Economics, & Organization* is currently published by Oxford University Press.

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/oup.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

---

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).



# Enforcing Rules on Oneself

THOMAS C. SCHELLING

*Harvard University*

## 1. INTRODUCTION

How does one devise a contract that it is easily enforced on each party by the other? How are rules designed for the behavior of each party, and incentives attached to compliance, so that the temptations and opportunities for non-compliance are minimized? How does an agreement make room for exceptions, for discretionary judgment about what is required, and for penalties on violations so that the whole arrangement need not collapse upon the first failure? How are the terms of agreement structured so that invisible noncompliance can be made visible or noncompliant intention be revealed in time to be challenged.

For contractual arrangements between two parties, this can be formulated as the question, How to compose the "bargain" to enhance the likelihood that it will be kept? The reciprocal bargain can be decomposed into the rules that each devises for the other: how do I devise your part of the bargain, the rules that you accept that I impose on you, to make it most likely that you will keep your part of the bargain? You have the same criteria for rules you impose on me, as part of the bargain. Often the reciprocal rules and incentives are interdependent: the sanction on my nonperformance is your release from a corresponding obligation, either some symmetrical obligation or any other part of the bargain that we have linked for purposes of incentives or equitable performance. But the rules we impose on each other are often quite asymmetrical. If I hire you to work for me, or lend you money, or rent you my home, there is typically no service you are to provide for me. If I do not want you to drink on the job or smoke in the presence of customers, the "rules" that I impose on your performance have no counterpart in the "rules" that you impose on me, such as that I pay you every week with a check that does not bounce.

I am interested in the special case of rules that people impose on themselves. I am thinking of people who wish not to oversleep, not to lose their

temper, not to drive after drinking too much, not to eat so much that they become overweight, not to abuse their children, not to violate the terms of parole, not to be cowardly, not to commit suicide. I am also thinking of people who wish to determine performance, not abstinence, such as pursuing some self-improvement regime like exercise or study, who wish not to procrastinate but to prepare their lectures and write their papers on time, or who wish to perform some singular feat such as confessing, enlisting, jumping by parachute, or committing suicide.

The question is how people devise effective rules to govern their future behavior, whether the behavior is continuous, regular or periodic, or single-shot; whether the behavior is abstinence or performance; whether the behavior is alone or as a member of a team with common interests; whether it is solitary or with someone's help. People resolve on courses of action and abstinence that they are apprehensive they may not fulfill; they want to constrain their own behavior at future moments in time when their preferences (or whatever impulses, temptations, phobias, fears, and passions control their choices) determine acts that are different from what they *now* prefer to do *then*. Sometimes the structure of the situation makes it possible to anticipate with finality that future decision—to burn the bridge today that one would use for escape tomorrow, to flush down the toilet the sleeping pills with which one might impetuously commit suicide, to throw the car keys into the darkness to keep from driving in search of sex, drugs, alcohol, tobacco, or the person one wishes to kill.

But that kind of anticipatory decisiveness is frequently unavailable. You cannot flush your fingernails down the toilet to keep from biting them, blind yourself to avoid wasting time at the television, or remain perpetually more than fifteen minutes from where cigarettes can be purchased. If you cannot preclude the behavior you deprecate or compel the behavior you hope for, you have to think about constraining that future behavior by means other than physical.

Legal means are not easy to come by. On another occasion (1984a) I looked at the question, Why can we not make legally enforceable vows—promises oriented toward no one in particular, intended to commit ourselves to the behavior we now wish to demand of ourselves? Why is it that certain constitutional rights are inalienable and we cannot contract to be held against our will until we have killed a drug habit, lost thirty pounds, or survived the full moon during which we acquire an urge to commit heinous crimes? Contracts are enforceable, subject to certain safeguards—not under duress, not between minors, not for involuntary servitude; promises are tantalizingly unenforceable unless they can be construed as part of an exchange. I cannot make a will that I cannot change; and it requires extraordinary ingenuity—is usually impossible—to get the state to tell me to do what I wish to be compelled to do, to intercede physically or to sanction with penalties the behavior that I now wish to avoid in the future.

We have, then, a territory in which "private ordering" is about all there is. We must devise rules for our own behavior that entail little or no reliance on the courts, not because we prefer something less expensive than recourse to the courts, prefer amicable rather than litigious relationships, prefer the greater certainty and greater promptness of resolution outside the courts, but because the courts refuse to extend us their jurisdiction. We cannot acquire standing; they disclaim the rules we want enforced.

So we have to make our own private arrangement. Mostly we depend on ingenuity, or creativity, or wisdom, in devising the rules we impose on ourselves and the incentives we can bring to bear. Sometimes we can lean on partners and teammates who share our interest in self-management, sometimes sharing a collective interest as when collective performance depends on our all showing up on time and being in shape; and sometimes we lean on a friend or kinsman who will accept the responsibility, even though there is no immediate reciprocity, someone who will hold our car keys or our wallet.

What I am going to do is look at some of the criteria for devising rules, the characteristics of rules that make them flimsy or robust, easily subverted or impregnable. I include the rules for rewards and penalties, the rules for exceptions to the rules, and the enlistment of help or technology in the surveillance and in the application of incentives. Without the help of someone else, rewards and penalties have to be self-administered; rules are needed to govern the administration not only of rewards and penalties for good and bad behavior but perhaps even of rewards and penalties for complying or failing to comply with the rules for administering the rewards and penalties themselves.

But before we consider the rules, some of the salient characteristics of the behaviors themselves ought to be noted. The kinds of behaviors for which we try to impose rules on ourselves differ drastically from each other, from rules against suicide to rules against picking scabs, from rules against losing our temper to rules against spending money foolishly, from rules about getting up in the morning to rules about fastening seatbelts, from rules about keeping things in order on one's desk to rules about writing a conference paper in plenty of time to revise it before the event and to get it into the hands of discussants at a decent interval before they are to discuss it.

I catalogued a few of these characteristics in an earlier paper (1984b) and will mention only a few here to suggest the kinds of things that may be pertinent. I have already mentioned the distinction between things that one must continuously avoid—picking fingernails, bad posture, wasteful day-dreaming; things that one must avoid on scheduled occasions, such as losing one's temper; things that are periodic, such as periodontal care or doing one's back exercises. Behaviors differ in their visibility: my guardian may need to see whether I sneak a quick drink in the kitchen; he can more readily check what I do on Friday night than watch me twenty-four hours a day, seven days a week. Some behaviors have the characteristic that if I fall behind I can make it

up, for example, running in the evening if I didn't run before breakfast, and sometimes there is a compensatory antidote, such as running an extra couple of miles if I ate an extra chocolate chip cookie. But there is no way to unsmoke a cigarette or to recover the hour that I wasted at television. Behaviors differ in elapsed time from moral lapse to misbehavior, cigarettes being typically more available than alcohol, alcohol more available than a gun, when one's resolve gives way. Some activities against which we try to apply constraints occur entirely within our minds—daydreaming, insomnia, depressing recollections, or captivation by an intellectual puzzle; most behaviors involve bodily activity of some sort.

I should take a moment to consider what our interest is. A possible interest is curiosity, ordinary or professional. People do impose rules on themselves, and there may be puzzling self-imposed constraints that are comprehensible when perceived as attempts at self-discipline. There may be some market transactions to understand: professional therapists and commercial consultants dispense advice about self-governing rules. We may have friends, colleagues, children, or spouses who ask our help, or appear to need help, in bringing their own behavior under better control, and we may even have friends or colleagues, children or spouses who appear misguided in the constraints they try to impose on themselves or have already, to our regret, succeeded in adopting. We may be interested in devising institutional arrangements that facilitate voluntary imposition of rules on oneself for certain kinds of behavior. We may want to discriminate between what we think are legitimate and illegitimate objectives, salutary and deleterious efforts at self-management, benign and dangerous technologies. And not least, maybe greatest, many of us have personal demands that will be satisfied by enhanced efficacy in command over our own behavior.

I offer the personal judgment that, by and large, people are more in need of greater efficacy in devising rules for their own behavior than in danger of shortsighted self-binding activity. So if a new discipline concerned with self-discipline became a source of effective advice to the people who want help, or to the therapists they go to for help, I would welcome rather than deplore the development. But like many welcome medications or therapies, techniques of effective self-denial and self-compulsion can be abused.

They can be abused by people whom we might consider immature or shortsighted or imprudent. The induction of a phobia, for example, may be safe and helpful in quitting intravenously administered drugs, even cigarettes, but dangerous in connection with eating, as anorexia nervosa testifies. People may invoke rules of avoidance—for example, with respect to the company they keep—that disconnect them from the very information that might rescue them from a restriction that they have outgrown.

And people may be coerced into severe "voluntary" restrictions on their own behavior. Youngsters may do something on a dare or to show off; momentary depression may induce radical efforts to restructure one's life.

Unbreachable rules may be accepted as a condition of employment, to one's detriment but perhaps not to one's employer's, if unbreachable rules could be had—refusal being not allowed as an option.

Generally I think that such dangers as there may be in providing people with mechanisms to precommit their own future behavior attach not so much to self-imposed *rules* as to the long-term or permanent technologies and sanctions and privations and disablements that people might choose to incur. Somebody may adopt rules of behavior in the interest of sexual continence that we find gratuitously ascetic or shortsighted, but compared with castration the decision is reversible, and if shortsighted the rules may not survive anyhow. Frustratingly difficult it often is to find rules that one can adopt effectively; the dangers that lurk are probably mostly confined to the neurotic and compulsive disorders.

### I. RATIONAL CHOICE AND ANTICIPATORY SELF-COMMAND

I should continue these preliminaries by putting our subject in the context of rational and efficacious behavior. I want to visualize the subject of these self-imposed rules as being thoughtfully and rationally engaged in devising constraints or influences on his own future behavior. But why does he have to decide now? And why, having made an anticipatory decision, can he not merely *remember* the decision rather than try to *impose* the decision in advance?

Some cases cause no theoretical perplexity. An absent-minded person who needs to remember to do an errand on the way to work may choose to drive an unaccustomed route, knowing that if he drives his usual route he will pursue his usual thoughts and nothing will remind him along the way of the errand he must stop for, while an unaccustomed route will continually remind him that there is some reason for his being in strange surroundings. Or if some rule is to govern behavior under stress one may need to decide on the rule in advance and practice it to become sufficiently habituated to abide by the rule when attention is elsewhere. The rule to count to ten before pulling a parachute ripcord may be a good way not only to measure the time needed to get clear of the plane but to discipline oneself against premature release. These examples remind us that even a paragon of rationality needs to anticipate episodes when time and attention for decision may be lacking and impulses contrary to prior intent may have to be contained.

The more anomalous cases are those in which our rational planner foresees an episode in which the preferences then governing his decisions—the actual choices he will make on that future occasion—are contrary to his current preferences with respect to that future occasion. Acting in accordance with his *current* preferences for that *future* occasion, he attempts to bind his behavior against those preferences that may then be governing. (In most common cases those alternate preferences come and go, haphazardly or predictably or

periodically; in some cases they may be long-lasting or permanent, deprecated and resisted nevertheless.) These are the cases that are harder to square with our usual image of the rationally deciding individual, because by convention, if not by scientific conviction, we typically impute continuity and stability to the individual and to the values that govern his preferences.

That does not mean that the epitome of rationality displays anything like unvarying appetite. The rational individual is assumed to recognize not only that his tastes may change progressively over time but that what he wants *now* depends on when *now* is. A person who is in no mood for comedy does not want anything frivolous right now but knows that when he has gotten over the blues he may enjoy and desire what would irritate now, just as people who do not now want their work interrupted by music can look forward to music this evening, and people who could not stomach a martini before breakfast can remember to replenish their gin supply so that they won't have to do without before dinner. The rational person is assumed to want his preferences to be satisfied; and the preferences that he wants satisfied are those in effect at the time that the "satisfaction" is to occur. There is nothing wrong with planning now to retire late in life to a part of the country that during the height of one's career would look like exile.

Thus the phenomenon we are examining is not merely the anticipation of changing preferences. It is *deprecating* expected or foreseeable changes, attempting to frustrate them rather than to accommodate them.

There is no archetype, no single theoretical model that captures some unique feature that epitomizes these situations. There are cases that have traditionally been known as "weakness of will," of temptation, of succumbing knowingly to what one is aware of earlier wishing to resist, even knowing that one will regret it. There is also expecting to succumb to anger, sympathy, or sexual passion with a powerful feeling of conviction, not of conflict. There is anticipated misbehavior under the influence of alcohol, medication, or fatigue. There are anticipated impulsive and compulsive behaviors in panic, phobia, or pain. And those are not all.

Most of these familiar efforts at anticipatory self-command are easier to recognize than to describe in theoretical terms, and this is not the occasion for attempting the theoretical identification. To most of us the phenomenon is familiar, whether or not we are comfortable with the way we construe it and whether or not we agree on its scientific description or its most fruitful conceptualization.

## 2. RULES, TACTICS, AND ADVICE

What kinds of rules work and what kinds do not? What kinds of sanctions lend themselves to self-imposed rules? What kinds of rules benefit from some friendly help and how can rules take advantage of technology?

The way I plan to pursue this question is by developing a classification. We

shall look at different kinds of rules, arranged in something like a hierarchy. But first a word about my use of the term rules. I mean any kind of command that one might like to issue to govern future behavior. It could be a complex of rules about how to behave toward one's children or how to conduct oneself in a stressful environment, for example, a meeting among one's peers. It could be a rule as simple as "walk, do not run," or a rule to get promptly out of bed when the alarm goes off. It could be a quantitative rule governing daily calorie intake or a flat prohibition such as no alcohol or no television. It could be a rule to cover rare contingencies (not losing one's temper with a police officer) or for periods of duress (never drinking unboiled water in India no matter how desperate with thirst), or a simple command performance (jumping out of the aircraft when the green light goes on).

The word *rule* has distinct meanings and there are correspondingly distinct devices in the management of one's own behavior. There are rules for assembling a child's tricycle or making an orange soufflé; there are rules about declaring taxable income, switching lanes in a tunnel, and plagiarizing. The latter are regulations to protect interests other than your own, and enforceability is pertinent; the former are only helpful guides, and you may follow them or not as you please, but it may be wise to follow them.

Something like the same distinction occurs in the rules of self-management. It is a good "rule" to avoid encounters with people who provoke you to lose your temper and misbehave; it is a good "rule" for binge eaters and binge shoppers to do their food shopping after breakfast when appetite is least active; it is a good "rule" to make exercise dates with good friends. But also it may be a good rule to count to ten before responding in anger, to lay your fork down between bites, to do your periodontic cleansing on a rigid schedule, or to clasp your hands firmly for three minutes if you have an urge to scratch your hives.

The former are like the instructions for assembling a tricycle: they are good advice. The latter require something like discipline, enforcement, or resistance to temptation.

The former are also what we might call *tactics*, which may be described in *rules* when they are frequently or regularly applicable. If Wednesday morning, sober and with no desire for alcohol, I anticipate that Friday night I may drink too much and become mean to my family, I may make a date to visit the country home of friends who do not drink and to spend the night; that is what I would call a "tactic". It is an efficacious maneuver undertaken when I am not at risk, when no particular discipline or act of will is required, to avoid putting myself at risk or to avoid succumbing when I am at risk. It would be a good *rule* to do that every Wednesday if indeed it is something that I can do every Wednesday. Such a rule would be only the regular or periodic exploitation of a tactic, and no problem of self-enforcement arises, not on Wednesday, nor on Friday if I am safely among teetotaling friends in a remote location by the time the urge to drink afflicts me. (Even drinking friends will take care of the



problem if my interest is solely in not being a menace to my family, and the drinking location is remote.)

I wish there were a better term than *tactic*. The vocabulary of efficacious conducts is unsatisfactory, *technique* being close in one of its senses, "a method of accomplishing a desired aim." Other terms are maneuver, expedient, ploy, artifice, strategem—most of which suggest trickery. Even *tactic* has the right meaning only in my 1937 *Webster's Collegiate*, where the plural is defined as "adroit devices for accomplishing an end."

My subject in this paper is neither the individual tactics that one may adroitly use to channel, constrain, or influence his own behavior on that future occasion when his voluntary choices would violate his current preferences nor the rules that merely prescribe or remind one of such tactics. My subject is the rules that are intended to restrict or to motivate behavior on those occasions and to exercise that influence *simply by virtue of being rules*. They are the rules that are intended to bind one against giving expression to those preferences that, in the absence of the rules, would govern his choice on that future occasion, and that do it simply by some moral authority or psychological force. They do it because they are invested with some recognizable authority. They represent a survival of the earlier values and preferences, a continuity, a link, a value that outlasts the preferences that motivated it.

The phenomenon exists. There are people, maybe everyone, who feel some obligation to perform as they solemnly resolved to perform or to abstain as they solemnly resolved to abstain even though they regret having incurred the obligation and would perform otherwise, or would fail to abstain, but for some nagging commitment that was motivated by values that at the moment have no claim or appeal. There is an analogy, probably more than mere analogy, with the force of a promise incurred earlier than one now regrets, even somewhat discredits, a promise that one could disregard without being caught, but not without espying himself in the act of breaking a promise.

Now that I have distinguished these two meanings of "rule" and identified my topic as the rules that appear to require discipline or enforcement, I want to back up a little. What I called a tactic, an adroit device for accomplishing an end, may be the adroit designing of a rule so that it is more easily enforced on oneself or requires less energy in enforcement. A good rule (in the one sense) may be to choose or to design rules (in the other sense) that are most readily, most easily, most reliably enforced. A good rule, if you must sit before a bowl of glistening salted peanuts during an hour's conversation, peanuts being what you are trying to eat as few of as possible, is not to eat the first peanut. Even if your diet allows a few peanuts, it may be easier to draw the line at zero than to stop at your quota after you have whipped your appetite into a frenzy with the first few peanuts. Knowing that the taste of peanuts excites the appetite further can help you design a rule—designing the rule can occur when there are no peanuts present and you are not even hungry—that is more readily enforced than a rule to stop after some number of peanuts. There is still a

problem of enforcing the rule you have chosen, but you have tactically chosen the most promising rule.

There are times when a rule, such as not looking in the direction of the peanut bowl, has to be taken in the presence of the temptation, while one's preference may be unsettled and vying for dominance. Someone who has resolved not to drink may be surprised by the announcement of a friend's daughter's engagement and have a glass of champagne thrust into her hand, or a dieter may be offered a pastry that his host's mother, present and smiling, spent the afternoon concocting in the kitchen. But mostly in this essay I shall be thinking of the rules that one devises for the occasion that has not yet begun in which one's preferences may shift drastically—the coldblooded determination of rules for the hotblooded encounter.

Other terms like *constraints* or *commands* would do as well as *rules*. And I include incentives and sanctions, rewards and penalties, as part of the system of rules or constraints or commands, keeping in mind that these are either self-administered (in accordance with other rules) or else informally arranged, without help from the legal system, through some sanction or authority or coercive capability that one arranges to invest in somebody helpful.

And the subject includes taking advantage of technology. Technology can be a substitute for rules, if one implants a device to deliver acute pain or a disabling chemical at the first detection of alcohol in the blood. It can be part of the rule system if one determines to ingest every day at noon a substance (antabuse) that will make alcohol cause sickness anytime within a dozen hours. That implant would be part of rule enforcement if it emitted a signal electronically to one's spouse or guardian. It is a little arbitrary to propose that an implanted device that delivers automatically a painful shock, and one that alerts a guardian who will administer a painful shock, are so different that one is a *substitute* for a rule and the other is an *appurtenance* to a rule; I don't think it matters whether I can draw the line. I just want to avoid recourse here to purely technological solutions—wiring jaws shut, implanting a year's supply of antabuse in a subcutaneous dispenser—while recognizing that technology may contribute to the monitoring of rules and even to the delivery of rewards and penalties.

### 3. GOALS, RULES, AND SUPPORTING RULES

At last we come to the question, What are the characteristics of rules that make them effective or ineffective? And what is this classification system that I have in mind?

At the top of the pyramid, above and beyond the rules, are what we might call goals or preferences. The purpose of rules is to help us reach our goals or satisfy our preferences. My *goal* is to be healthy, good-looking, productive; to have friends and a good reputation; to enjoy my various appetites. A *rule* might be to avoid overweight while eating nutritious foods that have few

carcinogens and taste good. But that merely translates a goal into relevant categories of behavior and does not tell me how to behave. It might be better to define rules as constraints on behavior. What about "keeping my weight at 140 pounds"? This sounds more like a rule, although it does not tell me what to do and what not to do. We could formulate it in calories: eat 16,000 calories per week. That I would call a rule. It may not be an effective one; maybe the rule should be to reduce weekly calories by 1,000 whenever my weight goes above 142 and enjoy an extra couple of hundred calories in any week when my weight goes below 138. One issue here is how to formulate a rule focused on a goal; another is how to tell a rule from a goal. I prefer to treat a rule as something that describes *behavior*, not *outcomes*, but rules frequently entail a feedback relation: reduce activity when outcome is above the target, increase activity when outcome is below the target.

These rules expressing the way we *wish* ourselves to behave I shall call *primary rules*. These primary rules are the behaviors we want to abide by. If there were no problem of self-management, all we would need is primary rules.

If we find ourselves eating too much we cut down, and if we seem to be losing weight we enjoyably eat more. If we are drinking too much we cut down, and if we are becoming a little too boisterous we quiet down. If I resolve not to drive home from a New Year's party if I have more than four drinks in four hours, and I can confidently count on remembering how many drinks I've had over how many hours and not driving home if it is more than those four, I don't need anything other than a primary rule. If the rule I want to follow is not to commit suicide no matter how depressed I feel, and having resolved not to commit suicide no matter how depressed I feel I know I shall not commit suicide no matter how depressed I feel, I do not need supporting rules. Primary rules suffice where enforcement is no problem. My interest is rules where enforcement is a problem.

Assuming I want primary rules devised in such a way that I can abide by them, what are the best characteristics they can have? This question arises precisely because of the difficulty of voluntarily following rules we have made for ourselves, especially in the circumstances in which we expect malfeasance. If I expect that on Friday evening I shall get drunk and become violent and want to beat my children, how do I devise a rule that will keep me, when drunk, from beating my children? Maybe the answer is to not get drunk. But that is not the primary rule: I have no interest in staying sober on Friday, only in not beating my children. In a moment we shall move on to precautionary or anticipatory rules. But there is a legitimate question of how to formulate the primary rule so that it is maximally likely to be effective.

There is some advice here that one finds in literature, in one's own experience, in the experience of friends, that is consistent with reasoning. Rules are best observed if you can easily tell the difference between compliance and violation. It seems more likely that one will abide by a self-imposed

rule if it is perfectly clear whether or not the rule has been adhered to or violated. Ambiguity is a bad characteristic.

Legal scholars use the notion of "bright lines" (bright lines=absence of ambiguity). Just discussing things abstractly, without regard to any particular substantive rule: things are probably more readily distinguishable if they are *qualitative* rather than *quantitative*. It may be harder to devise a convincing limit on the alcohol content of what one drinks than to have a rule against drinking alone, drinking before sundown, or drinking when one is going to drive home. Sometimes this criterion can be stated as qualitative versus quantitative, as in never drinking before the evening meal, never smoking until after dinner, never drinking "hard liquor." I think the better notion is that of *continuous* compared with *discrete*. However you do it, a "bright line" is a discontinuity, or a qualitative difference. I know whether or not I drank, smoked, ate dessert, turned on the television when I got home, or got up when the alarm went off. More complicated rules make better sense but do not seem to be as compelling.

One of the best discrete, qualitative discontinuities is between nothing and something. Not smoking at all is different from smoking any number of cigarettes. Zero is a unique quantity. It won't work for calories, but for addictive behaviors it can.

Another application of the discrete rule is *no exceptions*. Doing your back exercises three or four times a week is easily subject to moral erosion; doing your back exercises every day without fail is a more sustainable discipline. Lapses are more quickly noticeable and the habit can be better divorced from rationalizing processes if you do it every day rather than a few days a week. (Whoever invented the seven-day week didn't realize how hard it was going to be to schedule self-discipline events, like every-other-day exercise, into that prime-numbered week.) Minimal reliance on discretion and minimizing the occasions for considering exceptions is a useful principle.

An idea that sounds gratuitous but has the authority of writers of self-management books behind it is "write it down." Two arguments recommend it.

One is precision: writing invites careful formulation of boundaries, exceptions, penalties, and rewards. The other is ceremony: formalizing the rule in writing, perhaps with witnesses, attaches moral authority and makes violation more threatening to one's integrity, raising the stakes. Some authors recommend the form of a contract, with oneself or with someone trusted (Kalish: 296).

When we deal with actions rather than abstinence, piecemeal schedules often help. People who are intimidated by completing a substantial manuscript can often face the need to write three pages in a day. Schools often prescribe papers of a length and seriousness that students never before have had to face. These schools have discovered that a sequential schedule of intermediate deadlines is helpful. Students must turn in first a prospectus,

then an outline, then a rough draft, then a progress statement, then a clean draft, then a response to comments on the first draft, then a second draft, and ultimately a final draft.

This is an interesting blend of two related but distinct principles. How do you devise rules that *you* can enforce on the students? And how can you devise rules that students can enforce on *themselves*?

Giving students a sequence of incremental deadlines makes it easier to impose sanctions for delinquency. At the same time, providing students a sequence of milestones makes it more likely that the students will manage their own behavior, meeting piecemeal intermediate goals, never falling fatally far behind, never facing a task that is more than incremental.

Those are some of the characteristics that primary rules should have if they are to be self-enforceable. Where do they come from? I hope most of them will appeal to your own experience. Some come from reputable texts on self-management (Azrin and Nunn; Hodgson and Miller; O'Leary and Wilson; Watson and Tharp). Some even come from animal experiments (Ainslie). I shall not go into detail, because you probably don't doubt the assertions. I can draw somewhat on the more established literature on enforceable agreements and contracts and simply rely on the translatability of some principles of more formal two-party reciprocal enforcements to the case of self-enforcement.

Next we come to the rules that are not primary. What are they? A generic term might be *supporting* rules, all the kinds of rules that reflect the inadequacy of primary rules to take care of our behavior.

A useful category is *precautionary* rules. A precautionary rule is a rule that, if abided by, keeps one from violating the primary rule. Why should one be interested in precautionary rules? Because they may be more readily enforced.

It is hard to get children not to wade too deep into the water; it is easier to tell them to stay out of the water. It is hard to teach a child to avoid bottles of poisonous substances; it is easier to teach a child to avoid the soap and bleach cabinet. It may be hard to enforce a rule against riotous behavior but the police can more readily enforce a rule against congregating in large groups. Groups of three may be a precautionary rule against "large groups," which is itself a precautionary rule against "potentially troublesome groups."

Generally, a precautionary rule is a more inclusive rule. It draws wider boundaries. It is more restrictive. Being more restrictive, it might seem harder to enforce: "do not turn on the television" compared with "turn it off immediately after the news"; "do not order dessert" compared with "eat only half of it"; "do not eat the foods that whet your desire for a cigarette", compared with "do not smoke." There appear to be several ways in which the more restrictive rule can be an easy one to abide by.

One possibility is simply the brighter line, the discreteness of the precautionary boundary. Compromise may be harder to manage than an absolute

rule; even a puppy can be more effectively trained to stay in the kitchen if the kitchen can be defined coterminously with the linoleum floor.

But precautionary rules can also help in avoiding the stimuli that would trigger or aggravate a problem in self-management. I mentioned avoiding the first peanut. It is easier for a compulsive gambler to stay away from Atlantic City than to vacation there without gambling, easier to stay outside a casino than to watch a crap game without playing. Not all problems in self-management entail avoidable stimuli, but many do.

*Precursor* behavior is pertinent here. There are many behaviors that one wishes to avoid that are triggered by other behaviors—precursors—that are not nearly so hard to avoid. To mention one that some may think trivial (but many will not): many behaviors involve faces, fingernails, scabs and pimples, whiskers, hair, cuticles, lips, teeth, and gums. For some twenty million Americans these “grooming compulsive behaviors” are a significant part, a serious part, of their lives and a major issue in self-management.

Often the following is true but not known to the victim. There is some innocuous exploratory activity, such as feeling one’s face, that is habitual but not addictive. It is done, but it could be easily quit. There is a response to this precursor activity, which is discovering a whisker or pimple or irregularity of some sort, resistance to which is exceedingly difficult, and that leads to an irresistible behavior of real consequence. It is often the case that when one learns to recognize the precursor it is easily avoided but the consequent is almost irresistible. An obvious strategy is avoiding the precursor, recognizing that the follow-on behavior is ungovernable. This is a principle of immense power.

Many people who invest anguish and discomfort in quitting cigarettes, and build their hopes on a campaign of quitting, relax when they have had something to drink and when they have relaxed may resume smoking. For some people, it’s easier to quit drinking than to quit smoking; the best safeguard against smoking relapse is not to drink when the sequel may be smoking.

The counterpart of precautionary rules is *reinforcing* rules. What are the rules that we may find easier to follow than the primary rules we believe in? I call these *reinforcing* rules. They have a particular affinity for affirmative rules compared with negative, that is, performance compared with abstinence. Many people who would like to exercise find exercise vastly more attractive if they buy splendid equipment or attractive attire. People prepay their dues at an exercise clinic so that they will be motivated to go regularly to a Nautilus or Universal gym.

Maybe we should distinguish reinforcing rules from reinforcing actions. It is a reinforcing *action* to prepay the cost of an exercise clinic or to buy a lovely set of golf clubs to induce one’s exercise. It is a reinforcing *rule* if it is somewhat enforceable and enhances the likelihood that one will abide by a

primary rule. An example is the rule to get up when the alarm goes off. Ordinarily, I must get up and run several miles. Getting up is not easy, and getting up to run several miles is even more forbidding. If I can devise a rule for just getting up, it is not inhibited by the need to go running; once up I may be able to run because the inhibition is no longer supported by still being in bed.

A *disabling* rule is one that, if enforced, precludes the possibility of violating the primary rule. Whereas the precautionary rule merely draws a brighter line a safer distance away from the activity enjoined by the primary rule, the disabling rule would put the prohibited activity altogether beyond reach. Keeping nothing to eat in the kitchen precludes a midnight snack; leaving your car at home and going by taxi guarantees that if you drink too much you won't be caught driving; cutting the electric plug off your television on Sunday morning will make it easier to write that conference paper on Sunday afternoon. Again, as with the precautionary rules, there has to be an expectation that though you might not resist violating the primary rule, you can somehow make yourself obey the disabling rule. How it works is that the temptation to misbehave becomes irresistible at predictable times—such as mealtimes, early morning hours, weekend evenings, when in particular company or engaging in certain activities, like eating, drinking, or driving an automobile—and the anticipatory disabling is done in advance while the alternative preferences are in control.

The positive counterpart to a disabling rule would be an *enabling* rule. Friends have told me that their productivity has increased strikingly since they invested several thousand dollars in a word processor, not mainly because the machine is more efficient but because it is more fun. Joining a team or finding a buddy in committing yourself to run, to play tennis, to go to the fitness club on a regular schedule, can enhance the likelihood and the frequency that you will run, play tennis, or exercise. You arrange the social commitment during the time when your motivation toward future exercise is at a peak, that is, when the exercise itself is not threatening, thus raising the incentive when the time to exercise arrives.

Next we need some rules to govern the exceptions to the rules. There are times when it is embarrassing to refuse a proffered drink or dessert, mornings when it is too cold and windy to run, and opportunities to play when according to your primary rule you should be working. The problem is that some of these excuses are within your control, or depend on judgment about how much somebody would be offended if you did not share in the champagne toast, or how windy it appears out the window on a dark winter's morning. Like the primary rules, the rules governing exceptions are least likely to promote demoralization if they are neat, simple, objectively defined, qualitative and discrete. Determining a specific exception well in advance—deciding to have one glass of champagne at the wedding reception or deciding not to go running Thanksgiving morning—so that the exception is granted at a time when it

exercises no attraction can provide confidence in the event that one is not merely succumbing to temptation. (The fact that, in the event, the temptation is present does not mean that you are succumbing to it if you have obtained prior authorization from yourself.)

There is even, in dealing with exceptions, the possibility of leaving it up to a referee. Most people do not have live referees that can be trusted to be impartial, but a chance mechanism can impose a sense of discipline, of abiding by a verdict, although in some fraction of the cases it will withhold permission. But virtually all self-imposed rules and judgments, and the tactics for constraining future choices, are nonoptimal compared with the best that one might wish for.

Of course, certain rules are not intended to admit any kind of exception. Any effort to avoid the temptation to suicide fails on the first exception. Rules against child or wife abuse are intended to be absolute. And it is hard to imagine a reformed heroin addict deliberately working out a schedule of exceptions for future occasions.

Similar to the granting of exceptions to the rule is the application of a rule that requires *discretionary judgment*. Whether it is too cold and windy to exercise this morning could be treated not as the qualifying circumstances for an exception to the rule of always running on Thursday, but the application of discretion to the rule always to run on Thursday unless it is too cold and windy. Not running on Thanksgiving may be clearly an exception, but not running if it is too hot and humid is discretion applied to a rule that depends on some judgment. A rule against drinking alone is not intended to send you down to the bar in a hotel, looking for a casual acquaintance with whom you can enjoy a drink. I used to have a rule against smoking before my evening meal; it was easy to take care of Christmas and Thanksgiving as being specified in the rule, but dinner on a westward airplane flight seemed to depend on whether I expected to eat again after arrival.

One way to soften the impact of exceptions is to offer a compensatory substitute. If you do not run in bad weather, your sense of self-discipline will probably be enhanced if you get up early nevertheless, rather than taking the bonus in further sleep.

Many people find that delays can work like safety valves. The rule that one can, at any time, vouchsafe himself an exception, even legitimate termination, offers relief from the oppressive sense of permanent separation from whatever it is that one is abstaining from. Going an hour or two, or twelve hours, without the substance may be tolerable when eternal denial has become unthinkable. One then grants, according to the appropriate rule for such exceptions, the right to smoke (or whatever) tomorrow, or tonight, at a specified hour. One can always wait that long, if waiting makes the difference between a self-respecting conformity with the rules and flagrant relapse. But part of the rule is that one can revoke the notice at any time; and once notice is revoked, it can be invoked again but not reinstated. That is, if one can always



smoke on four hours notice, and one gives himself notice, and changes his mind during the first four hours', he can repeat the four hours' notice but cannot count the time elapsed earlier. Therapists assert that this works; I've known people who allege that it works for them; it makes sense, of a kind.

Exceptions and discretionary judgment blend into a third category, *violations*. We need a rule for handling violations. Perhaps there cannot be a way to accommodate violation of the rule against abusing one's own children, and suicide is uniquely irrelevant here, but people will smoke or drink or eat or rely on tranquilizers or lose their temper or watch television when they should be working and, on the analogy of the rules we might impose on servants, employees, or people who do not fulfill the contract, we have to have a way of coping.

There are two separate issues here. One is how to make violation unattractive, through the attachment of credible penalties. The other is how to keep violation from causing the whole enterprise to collapse, putting the rule in abeyance, discipline shattered, the episode finished.

A possibility would be to attach the severest penalties for the absolute deterrence of misbehavior and the absolute compulsion of required performance. But extreme penalties are not good if violation is inevitable, as it is for certain kinds of behavior; and it is not necessarily best to treat any violation as the equivalent of a criminal offense.

Take calories: the only thing wrong with violating the strictures on eating too much is that if you do it today you are more likely to do it tomorrow. As far as health and looks are concerned, you can always compensate tomorrow or next week. With calories there is an easy currency in which to exact the penalty: overcompensate tomorrow.

With cigarettes, that will work only if you are cutting down, not quitting. A rule that you send a check for \$100 to the political campaign of your favorite candidate's opponent for every cigarette you smoke is a good way of making sure that you will never persuade yourself, after a good dinner with wine, that a single cigarette by itself cannot do any harm. Is the penalty credible? Would you believe that tomorrow you will actually send the \$100 check to a political candidate you despise? Keep in mind that tomorrow the best way to persuade yourself that you are still a nonsmoker, despite smoking last night, is to celebrate the continued existence of the rule against smoking by invoking the rule on violations, sending the check.

In the same way we may want rules regarding *rewards* for good behavior, even extra rewards for behavior beyond the call of duty. Books on behavior therapy suggest that, at least early in a regime of abstinence, rewards for a week's compliance be made quite lavish.

Rules can provide a role for other people. I mentioned the possibility of doing things in teams, with people who have similar rules and similar needs for joint enforcement. But people can also be referees and judges. By *referee* I

mean somebody who applies the discretion you need in applying the rule or seeking an exception. Even if the person is biased toward leniency, as a spouse or friend may be, having to request the favor may itself be something of a deterrent; and if the referee takes his or her role seriously because of an interest in your succeeding with what you are trying to do, you may be genuinely able to count on their professional decisions.

I refer to such a person as a person as a referee if he or she has only moral authority—the right to grant or withhold an exception, to make a discretionary judgment, or to determine that a violation has occurred and that the penalty is due, or even determine the penalty. By a *judge* I mean somebody who has authority or physical possession. You give somebody your wallet when you go shopping, your car keys when you arrive at the party.

In between would be the rule that your wife serves the drinks. Her discretion, in measuring out the alcohol, is similar to that of a referee; she could as well watch you pour the alcohol into the glass, nodding her head or raising her eyebrow. Letting her do the actual pouring is a little more like giving her custody, although custody of the liquor is not physical in this case. But then actual custody of your car keys at that party is probably moral rather than physical: you are reluctant to ask for them; the person you gave them to would resist persuasively but probably hand them over if you became abusive.

The final category that I want to mention is the rule for handling *break-down*. This is what is sometimes called relapse, sometimes recidivism. It differs qualitatively from “violation” in being a violation for which no allowance in the rules was made, threatening collapse of the original resolve. It arises most frequently in the violation of an absolute rule, a flat prohibition, not a quantitative or continuous rule about calories eaten, miles run, hours put in on the job, or more alcohol than is good for one. The smoker, alcoholic, wife beater, gambler, or heroin addict who has quit typically does not treat relapse as a momentary episode, one to be punished according to the rules, one that may serve as a lesson, and one from which a person promptly recovers and gets back on course.

There is a dilemma. Among the most powerful motives for not having a first cigarette after three months’ abstinence, or a first drink, is the recognition that the slightest relapse precipitates total collapse: the effort is over, the investment wasted, the whole attempt a failure. But when a person does have that cigarette, despite the dire warning, it is in his interest to believe that he has merely violated the law, not overthrown the constitution; that despite the violation he is still a law-abiding citizen, that one cigarette after two or three months is pretty good if not perfect, and that he should pick himself up and get back on course.

Learning how to resolve this dilemma, for a particular person and for a particular rule, is one of the greater unresolved dilemmas in self-management.

REFERENCES

- Ainslie, George. 1975. "Specious Reward: A Behavioral Theory of Impulsiveness and Impulse Control," 82 *Psychological Bulletin* 463-96.
- Azrin, Nathan H., and Nunn, R. Gregory. 1977. *Habit Control in a Day*. New York: Simon and Shuster.
- Hodgson, Ray, and Miller, Peter. 1982. *Self-Watching—Addictions, Habits, Compulsions: What to Do About Them*. New York: Facts on File.
- Kalish, Harry I. 1981. *From Behavioral Science to Behavior Modification*. New York: McGraw-Hill Book Co.
- O'Leary, K. Daniel, and Wilson, G. Terrence. 1975. *Behavior Theory: Application and Outcome*. Englewood Cliffs: Prentice Hall.
- Schelling, Thomas C. 1984a. *Choice and Consequence*. Cambridge: Harvard University Press.
- . 1984b. "Self-Command in Practice, in Policy, and in a Theory of Rational Choice," 74 *American Economic Review* 1 (May).
- Watson, David L., and Tharp, Roland. 1981. *Self-Directed Behavior: Self-Modification for Personal Adjustment*. Monterey: Brooks/Cole Publishing.